

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 16:59:33 ; Search time 30893 Seconds  
(without alignments)  
11720.704 Million cell updates/sec

Title: US-10-029-020-13  
Perfect score: 8354  
Sequence: 1 gtttggtggtgtggaggac.....ccggaggtgacagagaggac 8354

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	8354	100.0	8354	6	AX556500	Sequence
2	8225	98.5	8645	6	AX600210	Sequence
3	7779.2	93.1	8438	6	AX675551	Sequence
4	6975.4	83.5	8585	10	AB025413	Mus muscu
5	6601.2	79.0	9722	10	AF059485	Mus muscu
6	4422.8	52.9	9264	5	AB026980	Danio rer
7	4197.2	50.2	8993	9	HSM806812	Homo sapi
8	3680.8	44.1	8624	9	AB037723	Homo sapi
9	3544.6	42.4	8964	10	AB025412	Mus muscu
10	3502.2	41.9	8645	6	AX662355	Sequence
11	3473.8	41.6	8675	6	AX662353	Sequence
12	3422	41.0	8473	6	AX662357	Sequence
13	3398.8	40.7	8487	6	AX662359	Sequence
14	3397.8	40.7	8689	6	AX250067	Sequence
15	3397.8	40.7	8689	10	AF086607	Rattus no
16	3385	40.5	9826	6	AX250008	Sequence
17	3384.8	40.5	8797	6	AX250063	Sequence
18	3384.8	40.5	8797	6	AX250066	Sequence
19	3384.8	40.5	8797	10	AB025411	Mus muscu
20	3351.4	40.1	9729	6	AX250013	Sequence
21	3330.4	39.9	8373	10	AB025410	Mus muscu
22	3320	39.7	8118	5	GGA238613	Gallus ga
23	3280	39.3	7816	10	AF195418	Mus muscu
24	3258	39.0	8297	9	AF100772	Homo sapi
25	3219	38.5	3486	9	HSM806114	Homo sapi
26	3165.8	37.9	8816	5	AB026979	Danio rer
27	3158.4	37.8	8575	6	AX921803	Sequence
28	3137.4	37.6	7781	9	AB032953	Homo sapi
29	3070.8	36.8	8409	5	GGA279031	Gallus ga
30	3070.8	36.8	8409	6	AX250068	Sequence
31	3055.6	36.6	3320	9	AK127101	Homo sapi
32	3051.8	36.5	5583	10	AK122490	Mus muscu
33	2925.2	35.0	7713	9	HSM808325	Homo sapi
34	2794.6	33.5	3048	9	AK127705	Homo sapi
35	2622.2	31.4	5804	10	AK122513	Mus muscu
36	2533	30.3	5560	6	AX250065	Sequence
37	2463	29.5	5309	9	AB040888	Homo sapi
38	2437.6	29.2	2685	9	AK056531	Homo sapi
39	1708.8	20.5	3111	6	BD270887	Method of
40	1708.8	20.5	3111	6	AR270934	Sequence
41	1708.8	20.5	3111	6	BD140153	Utilizati
42	1708.8	20.5	3111	9	AF009227	Homo sapi
43	1673.6	20.0	1680	6	AR270935	Sequence
44	1618.8	19.4	170156	2	AP002412	Homo sapi
45	1618.8	19.4	186084	9	AP002768	Homo sapi

ALIGNMENTS

RESULT 1  
AX556500  
LOCUS AX556500  
DEFINITION Sequence 13 from Patent WO02057453.  
ACCESSION AX556500  
VERSION AX556500.1 GI:25899736  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,  
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,  
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,

Smithson,G., Li,L. and Ji,W.  
Polypeptides and nucleic acids encoding same  
Patent: WO 02057453-A 13 25-JUL-2002;  
Curagen Corporation (US)

FEATURES

Location/Qualifiers  
1..8354  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 8354; DB 6; Length 8354;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 8354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTTGTGGATGTGGAGGAGCGCGGCGGCGGAGGCCATGGACGTGAAGGAGAGGAGCCTTA 60  
|||  
Db 1 GTTTGTGGATGTGGAGGAGCGCGGCGGCGGAGGCCATGGACGTGAAGGAGAGGAGCCTTA 60

QY 61 CCGTCTGCTGACCGCGCGCGGCGGCGGAGCGCGCTACACGAGCTCGTCCGCGGACAG 120  
|||  
Db 61 CCGTCTGCTGACCGCGCGCGGCGGCGGAGCGCGCTACACGAGCTCGTCCGCGGACAG 120

QY 121 CGAGGAGGCAAGCCCGCAGAAATCGTACAGTCCAGCGAGACCCCTGAAGGCCTACGA 180  
|||  
Db 121 CGAGGAGGCAAGCCCGCAGAAATCGTACAGTCCAGCGAGACCCCTGAAGGCCTACGA 180

QY 181 CCAGGACGCCCGCTAGCCTATGGCAGCCGCGTCAAGGACATTGTGCCGAGGAGGCCGA 240  
|||  
Db 181 CCAGGACGCCCGCTAGCCTATGGCAGCCGCGTCAAGGACATTGTGCCGAGGAGGCCGA 240

QY 241 GGAATTCTCCGCGACAGGTGCCAACTTCACCCCTCGGGAGCTGGGGCTGGAAGAATAAC 300  
|||  
Db 241 GGAATTCTCCGCGACAGGTGCCAACTTCACCCCTCGGGAGCTGGGGCTGGAAGAATAAC 300

QY 301 GCCCCTCACGGGACCCCTGTACCGGACAGACATTTGGCCTGCCCAATGCGGCTACTCCAT 360  
|||  
Db 301 GCCCCTCACGGGACCCCTGTACCGGACAGACATTTGGCCTGCCCAATGCGGCTACTCCAT 360

QY 361 GGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTCCCTGAGCACCCCGT 420  
|||  
Db 361 GGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTCCCTGAGCACCCCGT 420

QY 421 GCGTCTGTGGGCGCGGAGCACACGGTCAAGGCGGAGCTCCTGCTGTCCAGCCGGGCCAA 480  
|||  
Db 421 GCGTCTGTGGGCGCGGAGCACACGGTCAAGGCGGAGCTCCTGCTGTCCAGCCGGGCCAA 480

QY 481 TTCCAATCTCACACTCACCGACACCGAGCATGAAACACTGAGACTCATCCGGGCGG 540  
|||  
Db 481 TTCCAATCTCACACTCACCGACACCGAGCATGAAACACTGAGACTCATCCGGGCGG 540

QY 541 CCTGCAGAACCAACGCGGCTCCGAGCGCCGCGCGCGCTCTCGACGCCCAACACCCC 600  
|||  
Db 541 CCTGCAGAACCAACGCGGCTCCGAGCGCCGCGCGCGCTCTCGACGCCCAACACCCC 600

QY 601 CAACGAGCACCAACGCGGCTCCATTAACTCCCTGAACCGGGGCAACTTCAACGCGGAG 660  
|||  
Db 601 CAACGAGCACCAACGCGGCTCCATTAACTCCCTGAACCGGGGCAACTTCAACGCGGAG 660

QY 661 CAACCCAGCCCGCCCGCAGGACCACTCGTCTCCGAGAGCCCGCTGCCGCGGCGC 720  
|||  
Db 661 CAACCCAGCCCGCCCGCAGGACCACTCGTCTCCGAGAGCCCGCTGCCGCGGCGC 720

QY 721 CCAGGAGCCTGCCACGCCCGCAGGAGAACTGGTGTCTCAACAGCAACATCCCTGGAGAC 780  
|||  
Db 721 CCAGGAGCCTGCCACGCCCGCAGGAGAACTGGTGTCTCAACAGCAACATCCCTGGAGAC 780

QY 781 CAGGAACCTAGGCAAGCAGCCATTCTTAGGACATTCAGGACCACTCATTTAGATGGA 840  
|||  
Db 781 CAGGAACCTAGGCAAGCAGCCATTCTTAGGACATTCAGGACCACTCATTTAGATGGA 840

QY 841 CATTCTCGGCGCTCCGCGCATGATGGGCTTACAGTGACGGGCACTTCTCTTAAGCC 900  
|||  
Db 841 CATTCTCGGCGCTCCGCGCATGATGGGCTTACAGTGACGGGCACTTCTCTTAAGCC 900

Db 841 CATTCTCGGCGCTCCGCGCATGATGGGCTTACAGTGACGGCACTTCTCTTCAAGCC 900

QY 901 TGGAGGCACCTCCCGCTCTTCTGCACCACATCACAGGGTACCCACTGACCTCCAGCAC 960  
|||  
Db 901 TGGAGGCACCTCCCGCTCTTCTGCACCACATCACAGGGTACCCACTGACCTCCAGCAC 960

QY 961 AGTGTACTCTCTCCGCGCGGACCCCTGCCCGCAGCACCTTCCGCGCGCGCTTAA 1020  
|||  
Db 961 AGTGTACTCTCTCCGCGCGGACCCCTGCCCGCAGCACCTTCCGCGCGCGCTTAA 1020

QY 1021 CCTCAAGAGCCCTCCAACTGTAAGTGAAGTGCAGCCCTGAGCGCCATCGTCAT 1080  
|||  
Db 1021 CCTCAAGAGCCCTCCAACTGTAAGTGAAGTGCAGCCCTGAGCGCCATCGTCAT 1080

QY 1081 CTCAGCCACTCTGGTCTCTCTGGCATACTTTGTGGCCATGCACCTTTGGCCTAAA 1140  
|||  
Db 1081 CTCAGCCACTCTGGTCTCTCTGGCATACTTTGTGGCCATGCACCTTTGGCCTAAA 1140

QY 1141 CTGGCACCTGCAGCCGATGGAGGGCAGATGTATGAGATCACGGAGGACACAGCCAGCAG 1200  
|||  
Db 1141 CTGGCACCTGCAGCCGATGGAGGGCAGATGTATGAGATCACGGAGGACACAGCCAGCAG 1200

QY 1201 TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGCACTGGCTTAGAGACCCC 1260  
|||  
Db 1201 TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGCACTGGCTTAGAGACCCC 1260

QY 1261 TGACAGGAAAGGCAAGAACACAGAGAGGAAAGCCAGTAGTTTCTTCCAGAGGACAG 1320  
|||  
Db 1261 TGACAGGAAAGGCAAGAACACAGAGAGGAAAGCCAGTAGTTTCTTCCAGAGGACAG 1320

QY 1321 TTTTATAGATTCTGGAGAAATTTGATGTGGAGGCGAGCCTCCAGAGATTCTCTCTGG 1380  
|||  
Db 1321 TTTTATAGATTCTGGAGAAATTTGATGTGGAGGCGAGCCTCCAGAGATTCTCTCTGG 1380

QY 1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCATCTGAAATCAATGTGC 1440  
|||  
Db 1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCATCTGAAATCAATGTGC 1440

QY 1441 TCTGGAAAGGCAGCCCTGGTTGGCAATTTATGGCAGAAAGGCTCCCTCTCACATAC 1500  
|||  
Db 1441 TCTGGAAAGGCAGCCCTGGTTGGCAATTTATGGCAGAAAGGCTCCCTCTCACATAC 1500

QY 1501 ACAGTTTGACTTTGTGAGCTGTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAG 1560  
|||  
Db 1501 ACAGTTTGACTTTGTGAGCTGTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAG 1560

QY 1561 CCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTCCCCCTCCAGCCATGAGACAGG 1620  
|||  
Db 1561 CCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTCCCCCTCCAGCCATGAGACAGG 1620

QY 1621 CTTTCATCCAGTATTTGGATTTCAGGAATCTGCACTTGGCTTTTACAAATGACGAAAGGA 1680  
|||  
Db 1621 CTTTCATCCAGTATTTGGATTTCAGGAATCTGCACTTGGCTTTTACAAATGACGAAAGGA 1680

QY 1681 GTCAGAAAGTGGTTTCTTCTCACCACTGCCATTGAGTCGGTGGATACTGCCCCAGCAA 1740  
|||  
Db 1681 GTCAGAAAGTGGTTTCTTCTCACCACTGCCATTGAGTCGGTGGATACTGCCCCAGCAA 1740

QY 1741 CTGCTATGGCAATGGTACTGTCATCTCTGGGACCTGCCACTGCTTCTGGGTTTCTCTGGG 1800  
|||  
Db 1741 CTGCTATGGCAATGGTACTGTCATCTCTGGGACCTGCCACTGCTTCTGGGTTTCTCTGGG 1800

QY 1801 CCCCAGCTGTGGCAGAGCCCTCCTGCCCGCTGCTCTGTAGCGGAAATGGCCCAATACATGAA 1860  
|||  
Db 1801 CCCCAGCTGTGGCAGAGCCCTCCTGCCCGCTGCTCTGTAGCGGAAATGGCCCAATACATGAA 1860

QY 1861 AGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGCTGAGTGCATGTCGCCCAACCA 1920  
|||  
Db 1861 AGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGCTGAGTGCATGTCGCCCAACCA 1920

QY 1921 GTGTATCGATGTGGCCTGCAGCAACCTATGGCACCTGTGCATCACGGGCACTGTGCAATCTGCAA 1980  
|||  
Db 1921 GTGTATCGATGTGGCCTGCAGCAACCTATGGCACCTGTGCATCACGGGCACTGTGCAATCTGCAA 1980

QY	1981	CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCAACATGTTACAGG	2040
Dd	1981		
QY	2041	CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCAACATGTTACAGG	2040
Dd	2041		
QY	2041	CCGGGGTGTCTGCGTGAGAGGGCAATGCCATTGCTTTGTGGATGGGGAGGCACCAACTG	2100
Dd	2041		
QY	2101	CGAGACCCCGAGGGCCACATGCTTAGACCAAGTGTTCAGGCCACGGAACTTCTCTCCCGGA	2160
Dd	2101		
QY	2161	CACCGGGCTTTGCAGCTGTGACCCCAAGTGGACTGGACACGACTGTTCTATCGAGATCTG	2220
Dd	2161		
QY	2221	TGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGG	2280
Dd	2221		
QY	2281	CTGGATGGGGCAGCCTGCGACCAAGCGGCTGCCACCCGCGTGTGCCGAGCATGGGAC	2340
Dd	2281		
QY	2341	CTGCCCGGACGGAAGTCCGAGTGCAGCTGCGTGGTGAATGSGCAACACTGCACCATCGC	2400
Dd	2341		
QY	2401	TCACCTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGGAG	2460
Dd	2401		
QY	2461	ATGTACCTTAGACCTGAATGGTTGGCACTGCCCTGCGTGTGCCAGCTGGAGAGGCTGG	2520
Dd	2461		
QY	2521	CTGTGACACTTCCATGGAGACTGCCCTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGCTGG	2580
Dd	2521		
QY	2581	GGTGGACTGCATGGACCCCTGTGACATCATCCAGGAGACACAGTCCCTGTGCATATCAACCCGCTGTG	2640
Dd	2581		
QY	2641	CCTTGGCTCCCTAACCCCTCTGGACATCATCCAGGAGACACAGTCCCTGTGTACAGCA	2700
Dd	2641		
QY	2701	GAACCTACACTCCTTCTATGACCGCATCAAGTTCCCTGTTGGCAGGACACAGTCCCTGTGTACAGCA	2760
Dd	2701		
QY	2761	AATCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTATTCGTGGCCCAAGTGAT	2820
Dd	2761		
QY	2821	GACATCAGATGGAACCCCTGGTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTT	2880
Dd	2821		
QY	2881	TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC	2940
Dd	2881		
QY	2941	CATCATCCTGGGTTTCGAGCGGGCACCTTTCATCACACAGGAGCACACCCCTGTGGCTGCC	3000
Dd	2941		
QY	3001	ATGGGATCGCTTCTTGTATGGAAACCATCATCATGAGACATGAGGAGAATGAGATTCC	3060
Dd	3001		

QY	3061	CAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTCTCTCCATCCCCTGAC	3120
Dd	3061		
QY	3121	GTCTTCGCCAGCTCTGTGCAGAGAAAGGCCCCCATTTGTGCCGGAATTCAGGCTTTGCA	3180
Dd	3121		
QY	3181	GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCC	3240
Dd	3181		
QY	3241	TGGCTACAAATCTCTCTGAGGATCAGCCCTCACCCACCCGACCATCCCTTCAACCTCAT	3300
Dd	3241		
QY	3301	GAAGGTGCACCTCATGCTAGCGGTGGAGGGCCGCTCTTCAGGAAGTGGTTCGCTGCAGC	3360
Dd	3301		
QY	3361	CCCAGACCTGTCTCTATTATTTTCAATTTGGGACAAAGACAGACGTCTACAAACGAGGTGTT	3420
Dd	3361		
QY	3421	TGGGCTTTCAGAAAGCCTTTGTTTCCGTGGGTATGAATATGAATCTTGCCCCAGATCTAAT	3480
Dd	3421		
QY	3481	CCTGTGGGAAAAAGAAACAACAGTGTGAGGGCTATGAAAATGACGGTCCAAAGCTTGG	3540
Dd	3481		
QY	3541	AGGATGGAGCCTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCTCTGCACAAAGG	3600
Dd	3541		
QY	3601	GAATGGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATATGGCAA	3660
Dd	3601		
QY	3661	TGGCGCCGGAGAGCATCTCCTGCCCCAGCTGCAACGGCCTTGTCTGACGGCAACAAGCT	3720
Dd	3661		
QY	3721	CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGTGATTTCAA	3780
Dd	3721		
QY	3781	CTACATTAGAAAGATCTTCCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGAATAA	3840
Dd	3781		
QY	3841	AGATTTCAGACATAGTCAAGTCCAGCACAAATACTACCTGGCCACAGACCCCATGAG	3900
Dd	3841		
QY	3901	TGGGGCCGTCTTCCCTTTCTGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGT	3960
Dd	3901		
QY	3961	GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCGCT	4020
Dd	3961		
QY	4021	CCCCTTTGATGACACTCGCTGCGGGATGTTGGGAAGGCCACAGAAAGCCACACTCACCAA	4080
Dd	4021		
QY	4081	TCCCAGGGGTATTACAGTGGACAAGTTTGGCTGATCTACTTCGTGGATGGCACCATGAT	4140
Dd	4081		
QY	4141	CAGACGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCACATC	4200

Db 4141 CAGACGATCGATCAGAAATGGGATCATCTCCACCTGTCGGCTCTAATGATCTCACATC 4200  
Qy 4201 AGCCGGCCACTCAGCTGTGATCTGTCTCATGGATATTTCCAGGTAAGACTGGGCC 4260  
Db 4201 AGCCGGCCACTCAGCTGTGATCTGTCTCATGGATATTTCCAGGTAAGACTGGGCC 4260  
Qy 4261 CACAGACTTAGCCATCAACCCAAATGGACAACCTCACTTTATGTCTCTCGACAACAATGTGGT 4320  
Db 4261 CACAGACTTAGCCATCAACCCAAATGGACAACCTCACTTTATGTCTCTCGACAACAATGTGGT 4320  
Qy 4321 CCTGCAAAATCTCTGAAAACCAACAGGTGCGCAATTTGTCGGGGAGGCCCATGCACTGCCA 4380  
Db 4321 CCTGCAAAATCTCTGAAAACCAACAGGTGCGCAATTTGTCGGGGAGGCCCATGCACTGCCA 4380  
Qy 4381 GGTCCTGGCATTTGACCACTTCTGCTAAGCAAGGTGGCCATCCACGCAACCTGGAGTC 4440  
Db 4381 GGTCCTGGCATTTGACCACTTCTGCTAAGCAAGGTGGCCATCCACGCAACCTGGAGTC 4440  
Qy 4441 AGCCACCGCTTTGGCTGTTTACACAATGGGGTCCCTGATATTTGCTGAGACTGATGAA 4500  
Db 4441 AGCCACCGCTTTGGCTGTTTACACAATGGGGTCCCTGATATTTGCTGAGACTGATGAA 4500  
Qy 4501 AAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4560  
Db 4501 AAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4560  
Qy 4561 CCCAGTGGCTGTGACTGTAAGAAATGATGCCAACTGTGTGATTTTCTGGAGACGATGG 4620  
Db 4561 CCCAGTGGCTGTGACTGTAAGAAATGATGCCAACTGTGTGATTTTCTGGAGACGATGG 4620  
Qy 4621 TTATGCCAAGGATGCAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGA 4680  
Db 4621 TTATGCCAAGGATGCAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGA 4680  
Qy 4681 GCTTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTATCCGGAAGAACAGCCTTT 4740  
Db 4681 GCTTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTATCCGGAAGAACAGCCTTT 4740  
Qy 4741 CCTCAACACCCAGAACATGATGAGCTGTCTTACCAATTTGACCAAGGCTCTATCTGTT 4800  
Db 4741 CCTCAACACCCAGAACATGATGAGCTGTCTTACCAATTTGACCAAGGCTCTATCTGTT 4800  
Qy 4801 TGATPACACCGGCAAGCACTGTATACACCCAAAGCTGCCACAGGAGCTACCTGTACAA 4860  
Db 4801 TGATPACACCGGCAAGCACTGTATACACCCAAAGCTGCCACAGGAGCTACCTGTACAA 4860  
Qy 4861 CTTACCTACACTGGGACCGCGACATCACACTCATCACAGACAACAATGGCAACATGGT 4920  
Db 4861 CTTACCTACACTGGGACCGCGACATCACACTCATCACAGACAACAATGGCAACATGGT 4920  
Qy 4921 AAATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGT 4980  
Db 4921 AAATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGT 4980  
Qy 4981 GTACTGGGTGACCATGGGACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGACACGA 5040  
Db 4981 GTACTGGGTGACCATGGGACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGACACGA 5040  
Qy 5041 GTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCCAAAGCAATGAAA 5100  
Db 5041 GTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCCAAAGCAATGAAA 5100  
Qy 5101 CGGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCC 5160  
Db 5101 CGGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCC 5160  
Qy 5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGTGATGTCAGGTAGA 5220  
Db 5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGTGATGTCAGGTAGA 5220  
Qy 5221 GACCTCCAGCAAGGATGTACCAATAACCAACCTGTCTGCTCAGGCGCCTTCTA 5280

Db 5221 GACCTCCAGCAAGGATGTGTCAACATAACCAACCACTGTCTGCTCAGGCGCCTTCTA 5280  
Qy 5281 CACACTGCTCAAGAACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCTTGG 5340  
Db 5281 CACACTGCTCAAGAACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCTTGG 5340  
Qy 5341 GCTGCTGTGGCCAAACGGGATGGAGGTGGCGCTGCAGACTGAGCCCCACCTGCTGGCTGG 5400  
Db 5341 GCTGCTGTGGCCAAACGGGATGGAGGTGGCGCTGCAGACTGAGCCCCACCTGCTGGCTGG 5400  
Qy 5401 CACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACCGTGGCCATCGACAACGGCCTCAA 5460  
Db 5401 CACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACCGTGGCCATCGACAACGGCCTCAA 5460  
Qy 5461 CCTGCTGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGGCG 5520  
Db 5461 CCTGCTGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGGCG 5520  
Qy 5521 CCGGCTGCGGCTGACAAACCGAAATCTCTATCTCTGGACTTTGATCGCGTAACACGCAC 5580  
Db 5521 CCGGCTGCGGCTGACAAACCGAAATCTCTATCTCTGGACTTTGATCGCGTAACACGCAC 5580  
Qy 5581 AGAGAAGATCTATGATGACCAACCGAAATTCCTATCTCTGGACTTTGATCGCGTAACACGCAC 5640  
Db 5581 AGAGAAGATCTATGATGACCAACCGAAATTCCTATCTCTGGACTTTGATCGCGTAACACGCAC 5640  
Qy 5641 GCGGCCCAGCCTCTGTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCC 5700  
Db 5641 GCGGCCCAGCCTCTGTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCC 5700  
Qy 5701 TGGGGTTTACATTTGCTGGCATCCAGAGGGCATCATGTCTGAAAGAAATGGAATACGACCA 5760  
Db 5701 TGGGGTTTACATTTGCTGGCATCCAGAGGGCATCATGTCTGAAAGAAATGGAATACGACCA 5760  
Qy 5761 GCGGGCCGCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATACTT 5820  
Db 5761 GCGGGCCGCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATACTT 5820  
Qy 5821 AGAGAAGTCCATGGTGTCTACTACACAGCCAGAGGCTATATCTTTGAGTTGACAA 5880  
Db 5821 AGAGAAGTCCATGGTGTCTACTACACAGCCAGAGGCTATATCTTTGAGTTGACAA 5880  
Qy 5881 GAATGACCGCCTCTCTTGTGACGATGCCCAACGTGGCGCGGACACTAGAGACCAT 5940  
Db 5881 GAATGACCGCCTCTCTTGTGACGATGCCCAACGTGGCGCGGACACTAGAGACCAT 5940  
Qy 5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAT 6000  
Db 5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAT 6000  
Qy 6001 ACAGGACTTCACTGAGGATGGGACCTCTTCAACCTTCTACCTGGGCACTGGCCGCGAG 6060  
Db 6001 ACAGGACTTCACTGAGGATGGGACCTCTTCAACCTTCTACCTGGGCACTGGCCGCGAG 6060  
Qy 6061 GGTGATATACAAGTATGGCAAACTGTCAAAGTGGCAGAGACGCTCTATGACACCAACCA 6120  
Db 6061 GGTGATATACAAGTATGGCAAACTGTCAAAGTGGCAGAGACGCTCTATGACACCAACCA 6120  
Qy 6121 GGTGAGTTTCACTATGACGAGACGGCAGGATGTGAAAGACCATCAACCTACAGAATGA 6180  
Db 6121 GGTGAGTTTCACTATGACGAGACGGCAGGATGTGAAAGACCATCAACCTACAGAATGA 6180  
Qy 6181 GGGCTTCACTGACCATCCGCTACCGTACCGTGGGCCCTGATTGACCGACAGATCTT 6240  
Db 6181 GGGCTTCACTGACCATCCGCTACCGTACCGTGGGCCCTGATTGACCGACAGATCTT 6240  
Qy 6241 CCGCTTCACTGAGGAAGGCTGGTCAACCGCCGTTTGGACTACAACCTATGACAACAGCTT 6300  
Db 6241 CCGCTTCACTGAGGAAGGCTGGTCAACCGCCGTTTGGACTACAACCTATGACAACAGCTT 6300  
Qy 6301 CCGGCTGACCAAGGATGTACCAATAACCAACCTGTCTGCTCAGGCGCCTTCTA 6360  
Db 6301 CCGGCTGACCAAGGATGTACCAATAACCAACCTGTCTGCTCAGGCGCCTTCTA 6360





QY 1861 AGCAGATGCTTGTGCCACAGTGGCTGGAAGGGCGCTGAGTGCATGTGCCCAACCA 1920  
Dd  
QY 1931 AGCAGATGCTTGTGCCACAGTGGCTGGAAGGGCGCTGAGTGCATGTGCCCAACCA 1990  
Dd  
QY 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGATCAACGGGCACCTGATCTGCAA 1980  
Dd  
QY 1991 GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGATCAACGGGCACCTGATCTGCAA 2050  
QY 1981 CCCTGGCTACAAGGGCGAGAGCTGTGAGGCAAGTGGACTGCATGGACCCCAATGTTGAGG 2040  
Dd  
QY 2051 CCCTGGCTACAAGGGCGAGAGCTGTGAGGCAAGTGGACTGCATGGACCCCAATGTTGAGG 2110  
QY 2041 CCGGGTGTCTGCGTGAAGGCGAATGCCATGCTTTGTGGATGGGAGGCACCAACTG 2100  
Dd  
QY 2111 CCGGGTGTCTGCGTGAAGGCGAATGCCACTGCTCTGTGGATGGGAGGCACCAACTG 2170  
QY 2101 CGAGACCCCGAGGGCCACATGCTTAGACCAAGCTGACTGGACACGACTGTTCTATCGAGATCTG 2160  
Dd  
QY 2171 CGAGACCCCGAGGGCCACATGCTTAGACCAAGCTGTTGAGGCAAGTGGACTGCATGGACCCCACTTCCCTCCCGGA 2230  
QY 2161 CACCGGCTTTCAGCTGTGACCAAGCTGACTGGACACGACTGTTCTATCGAGATCTG 2220  
Dd  
QY 2231 CACCGGCTTTCAGCTGTGACCAAGCTGACTGGACACGACTGTTCTATCGAGATCTG 2290  
QY 2221 TGCTGCCGACTGTGGTGCCCATGGCGTGTGCGTAGGGGACCTGCCGCTGCCGAGGATGG 2280  
Dd  
QY 2291 TGCTGCCGACTGTGGTGCCCATGGCGTGTGCGTAGGGGACCTGCCGCTGCCGAGGATGG 2350  
QY 2281 CTGGATGGGGGACGCTGCCGACCAAGGGGCTGCCACCCCGCTGTGCCGAGCATGGGAC 2340  
Dd  
QY 2351 CTGGATGGGGGACGCTGCCGACCAAGGGGCTGCCACCCCGCTGTGCCGAGCATGGGAC 2410  
QY 2341 CTGCCGACCGGCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2400  
Dd  
QY 2401 TCACATCTGGATAGGATAGTTAAAGAGGGTGGCTTGGCTGGAGTGGAGTGGAGTGGAG 2460  
Dd  
QY 2471 TCACATCTGGATAGGATAGTTAAAGAGGGTGGCTTGGCTGGAGTGGAGTGGAGTGGAG 2530  
QY 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGTGGCTGGAGAGGAGCTGG 2520  
Dd  
QY 2531 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGTGGCTGGAGAGGAGCTGG 2590  
QY 2521 CTGTGACACTTCCATGGAGACTGCGTGGCTGACAGCAACAGCAATGATGGAGTGGCCT 2580  
Dd  
QY 2591 CTGTGACACTTCCATGGAGACTGCGTGGCTGACAGCAACAGCAATGATGGAGTGGCCT 2650  
QY 2581 GGTGACTGATGGACCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640  
Dd  
QY 2651 GGTGACTGATGGACCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2710  
QY 2641 CCTTGGCTCCCTAACCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGCA 2700  
Dd  
QY 2711 CCTTGGCTCCCTAACCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGCA 2770  
QY 2701 GAACCTACACTCCTTCTATGACCCGATCAAGTTCCCTCGTGGGAGGACAGCACGACAT 2760  
Dd  
QY 2771 GAACCTACACTCCTTCTATGACCCGATCAAGTTCCCTCGTGGGAGGACAGCACGACAT 2830  
QY 2761 AATCCCCGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTATTGCTGGCCAAAGTAT 2820  
Dd  
QY 2831 AATCCCCGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTATTGCTGGCCAAAGTAT 2890  
QY 2821 GACATCAGATGGAACCCCTGGTTGGTGAACATCAGTTTGTCAATAACCCCTCTCTT 2880  
Dd  
QY 2891 GACATCAGATGGAACCCCTGGTTGGTGAACATCAGTTTGTCAATAACCCCTCTCTT 2950  
QY 2881 TGGATATACAAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGATCTC 2940  
Dd  
QY 2951 TGGATATACAAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGATCTC 3010

QY 2941 CATCATCCTGCGGTTTCGAGCGGCGACCTTTTCATCACACAGGAGCACACCTGTGGTGCC 3000  
Dd  
QY 3011 CATCATCCTGCGGTTTCGAGCGGCGACCTTTTCATCACACAGGAGCACACCTGTGGTGCC 3070  
Dd  
QY 3001 ATGGGATCGCTTCTTTGTTCATGGAACCATCATCATGAGACATGAGGAGAAATGAGATTCC 3060  
Dd  
QY 3071 ATGGGATCGCTTCTTTGTTCATGGAACCATCATCATGAGACATGAGGAGAAATGAGATTCC 3130  
QY 3061 CAGCTGTGACCTGAGCAATTTTCCCGCCCAACCCAGTCGTCTCTCCATCCCCACTGAC 3120  
Dd  
QY 3131 CAGCTGTGACCTGAGCAATTTTCCCGCCCAACCCAGTCGTCTCTCCATCCCCACTGAC 3190  
QY 3121 GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCAATTTGTCGCGAAATTCAGGCTTTGCA 3180  
Dd  
QY 3191 GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCAATTTGTCGCGAAATTCAGGCTTTGCA 3250  
QY 3181 GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCC 3240  
Dd  
QY 3251 GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCC 3310  
QY 3241 TGGCTACAAATCTCTCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCC 3300  
Dd  
QY 3311 TGGCTACAAATCTCTCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCC 3370  
QY 3301 GAAGGTGCACCTCATGGTAGCGGTGGAGGGCGCCCTCTTCAGGAAGTGGTTCGCTGCAGC 3360  
Dd  
QY 3371 GAAGGTGCACCTCATGGTAGCGGTGGAGGGCGCCCTCTTCAGGAAGTGGTTCGCTGCAGC 3430  
QY 3361 CCCAGACCTGTCTTATTTTCATTTGGGACAAGACAGACGCTTACAAACCAAGGTTGT 3420  
Dd  
QY 3431 CCCAGACCTGTCTTATTTTCATTTGGGACAAGACAGACGCTTACAAACCAAGGTTGT 3490  
QY 3421 TGGGCTTTCAGAAAGCCTTTGTTTCCGTGGGTTATGAATATGAATTCCTGCCAGATCTAAT 3480  
Dd  
QY 3491 TGGGCTTTCAGAAAGCCTTTGTTTCCGTGGGTTATGAATATGAATTCCTGCCAGATCTAAT 3550  
QY 3481 CCTGTGGGAAAAAGAACCAACAGTGTCTGAGGGCTATGAAATTCAGCGCTTCAAGCTTGG 3540  
Dd  
QY 3551 CCTGTGGGAAAAAGAACCAACAGTGTCTGAGGGCTATGAAATTCAGCGCTTCAAGCTTGG 3610  
QY 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCCTGCACAAAGG 3600  
Dd  
QY 3611 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCCTGCACAAAGG 3670  
QY 3601 GAATGGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCCTGTTCATTTGGGAGCATCATGGCAA 3660  
Dd  
QY 3671 GAATGGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCCTGTTCATTTGGGAGCATCATGGCAA 3730  
QY 3661 TGGGCGCCGAGAAAGCATCTCCTGCCCCAGCTGCAACGGCCTTGTGACGGCAACAAGCT 3720  
Dd  
QY 3731 TGGGCGCCGAGAAAGCATCTCCTGCCCCAGCTGCAACGGCCTTGTGACGGCAACAAGCT 3790  
QY 3721 CCTGGCCCGAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA 3780  
Dd  
QY 3791 CCTGGCCCGAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA 3850  
QY 3781 CTACATTAGAGGATCTTCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGAATAA 3840  
Dd  
QY 3851 CTACATTAGAGGATCTTCCCTCTGGAATGTCAACCAACATCCTAGAGCTG----- 3902  
QY 3841 AGATTTCAGACATAGTCCAGTCCAGCACAAATACTA CCTGGCCACAGACCCCATGAG 3900  
Dd  
QY 3903 -----AGTCACAGTCCAGCACAAATACTACCTGGCCACAGACCCCATGAG 3949  
QY 3901 TGGGCGCGTCTTCCCTTTCAGACAGCAACAGCCGGGCTCTTTAAATCAAGTCCACTGT 3960  
Dd  
QY 3950 TGGGCGCGTCTTCCCTTTCAGACAGCAACAGCCGGGCTCTTTAAATCAAGTCCACTGT 4009  
QY 3961 GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAAGTGCCT 4020  
Dd  
QY 4010 GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAAGTGCCT 4069  
QY 4021 CCCCTTTGATGACACTCGCTCGGGGATGTTGGGAAGGCCACAGAGCCACACTCACCAA 4080

Db 4070 CCCCCCTTGTGATGACACTCGCTGGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACCAA 4129  
Qy 4081 TCCAGGGGTATTACAGTGGCAAGTTTGGGCTGATCTACTTCTGATGGCACCATGAT 4140  
Db 4130 TCCAGGGGCATTACAGTGGCAAGTTTGGGCTGATCTACTTCTGATGGCACCATGAT 4189  
Qy 4141 CAGACGCATCGATCAGAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC 4200  
Db 4190 CAGACGCATCGATCAGAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC 4249  
Qy 4201 AGCCCCGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCAGGTAAGACTGGAGTGGCC 4260  
Db 4250 AGCCCCGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCAGGTTTCCCTGGAGTGGCC 4309  
Qy 4261 CACAGACTTAGCCATCAACCCCAATGGACAACCTCAGTTATGTCCTCGACAACAATGTGGT 4320  
Db 4310 CACAGACTTAGCCATCAACCCCAATGGACAACCTCAGTTATGTCCTCGACAACAATGTGGT 4369  
Qy 4321 CCTGCAATCTCTGAAAAACCCAGGTCGCGCATTTGTCGCGGGAGGCCCATGCACTGCCA 4380  
Db 4370 CCTGCAATCTCTGAAAAACCCAGGTCGCGCATTTGTCGCGGGAGGCCCATGCACTGCCA 4429  
Qy 4381 GGTCCTGGCATTTGACCACTTCTGTAAGCAAGGTGGCCATCCACGCAACCTGGAGTC 4440  
Db 4430 GGTCCTGGCATTTGACCACTTCTGTAAGCAAGGTGGCCATCCACGCAACCTGGAGTC 4489  
Qy 4441 AGCCACCGCTTTGGCTGTTTACACAATGGGGTCTGTATATTTGCTGAGACTGATGAGAA 4500  
Db 4490 AGCCACCGCTTTGGCTGTTTACACAATGGGGTCTGTATATTTGCTGAGACTGATGAGAA 4549  
Qy 4501 AAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4560  
Db 4550 AAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4609  
Qy 4561 CCCAGTGGCTGTGACTGTAAAAAATGATGCCAATCTGATGTTTCTGGAGACGATGG 4620  
Db 4610 CCCAGTGGCTGTGACTGTAAAAAATGATGCCAATCTGATGTTTCTGGAGACGATGG 4669  
Qy 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGA 4680  
Db 4670 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGA 4729  
Qy 4681 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTT 4740  
Db 4730 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTT 4789  
Qy 4741 CCTCAACACCCAGAACATGATGAGCTGTCTTCAACCAATTGACCAAGAGCTCTATCTGTT 4800  
Db 4790 CCTCAACACCCAGAACATGATGAGCTGTCTTCAACCAATTGACCAAGAGCTCTATCTGTT 4849  
Qy 4801 TGATACCAACCGCAAGCCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAA 4850  
Db 4850 TGATACCAACCGCAAGCCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAA 4909  
Qy 4861 CTTACCTACACTGGGGACGGCGACATCACACTCATCAGACAACAATGGCAACATGGT 4920  
Db 4910 CTTACCTACACTGGGGACGGCGACATCACACTCATCAGACAACAATGGCAACATGGT 4969  
Qy 4921 AAATGTCCGCCGAGACTCTACTGGGATGCCCTTCTGGCTGGTGGTCCAGATGGCCAGGT 4980  
Db 4970 AAATGTCCGCCGAGACTCTACTGGGATGCCCTTCTGGCTGGTGGTCCAGATGGCCAGGT 5029  
Qy 4981 GTACTGGGTGACCATGGGCACCAACAGTGCATCTAAGAGTGTGACCAACAAGGACACGA 5040  
Db 5030 GTACTGGGTGACCATGGGCACCAACAGTGCATCTAAGAGTGTGACCAACAAGGACACGA 5089  
Qy 5041 GTTGGCCATGATGACATACCATGGCAATTCGGCTTCTGGCAACCAAAAGCAATGAAAA 5100  
Db 5090 GTTGGCCATGATGACATACCATGGCAATTCGGCTTCTGGCAACCAAAAGCAATGAAAA 5149  
Qy 5101 CGGATGGACAACATTTATGAGTACGACAGCTTTGGCCGCTGACAAATGTGACCTTCCC 5160

Db 5150 CGGATGGACAACATTTTATGAGTACGACAGCTTTTGGCCGCTGACAAATGTGACCTTCCC 5209  
Qy 5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCAGGTAGA 5220  
Db 5210 TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCAGGTAGA 5269  
Qy 5221 GACCTCCAGCAAGGATGATGTACCAATAACCAACCACTGTCTGCCTCAGGCGCTTCTA 5280  
Db 5270 GACCTCCAGCAAGGATGATGTACCAATAACCAACCACTGTCTGCCTCAGGCGCTTCTA 5329  
Qy 5281 CACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGATGGCTCTTGGC 5340  
Db 5330 CACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGATGGCTCTTGGC 5389  
Qy 5341 GCTGCTGCTGGCCAAACGGCATGGAGGTGGCGTGCAGACTGAGCCCCACTTGTGGCTGG 5400  
Db 5390 GCTGCTGCTGGCCAAACGGCATGGAGGTGGCGTGCAGACTGAGCCCCACTTGTGGCTGG 5449  
Qy 5401 CACCGTCAACCCACCCGTCGGGCAAGAGGAATGTCAAGCTGCCATCGACAACGGGCCTCAA 5460  
Db 5450 CACCGTCAACCCACCCGTCGGGCAAGAGGAATGTCAAGCTGCCATCGACAACGGGCCTCAA 5509  
Qy 5461 CCTGCTGAGTGGCGCCAGCGCAAGTTTCAACCTTGGATTCTGTACGACCAAGCGGG 5520  
Db 5510 CCTGCTGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGTCTTGGCG 5569  
Qy 5521 CCGGCTGCGGGTGCAACCGGAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCAC 5580  
Db 5570 CCGGCTGCGGGTGCAACCGGAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCAC 5629  
Qy 5581 AGAGAAGATCTATGATGACCAACCGCAAGTTTCAACCTTGGATTCTGTACGACCAAGCGGG 5640  
Db 5630 AGAGAAGATCTATGATGACCAACCGCAAGTTTCAACCTTGGATTCTGTACGACCAAGCGGG 5689  
Qy 5641 GCGGCCAGCCTCTGCTCAGCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC 5700  
Db 5690 GCGGCCAGCCTCTGCTCAGCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC 5749  
Qy 5701 TGGGGTTACATTGCTGGCATCCAGAGGGGCTCATGTCTGAAAGAATGGAATACGACCA 5760  
Db 5750 TGGGGTTACATTGCTGGCATCCAGAGGGGCTCATGTCTGAAAGAATGGAATACGACCA 5809  
Qy 5761 GCGGGCCGCATCACATCCAGGATCTTCCGCTGATGGGAAGACATGGAGCTACACATACTT 5820  
Db 5810 GCGGGCCGCATCACATCCAGGATCTTCCGCTGATGGGAAGACATGGAGCTACACATACTT 5869  
Qy 5821 AGAGAAGTCCATGGTGTCTACTACAGCCAGGCGAGTATATCTTTGAGTTGACAA 5880  
Db 5870 AGAGAAGTCCATGGTGTCTACTACAGCCAGGCGAGTATATCTTTGAGTTGACAA 5929  
Qy 5881 GAATGACCGCCTCTCTTCTGTGACGATGCCCAAGCTGGCGGAGACTAGAGACCAT 5940  
Db 5930 GAATGACCGCCTCTCTTCTGTGACGATGCCCAAGCTGGCGGAGACTAGAGACCAT 5989  
Qy 5941 CCGCTCAGTGGGCTACTACAGAAAACATCTATCAGCCCCCTGAGGGCAATGCCCTCAGTCA 6000  
Db 5990 CCGCTCAGTGGGCTACTACAGAAAACATCTATCAGCCCCCTGAGGGCAATGCCCTCAGTCA 6049  
Qy 6001 ACAGGACTTCACTGAGGATGGCACCTCTTCAACACCTTCTACCTGGGCACTGGCCGCGAG 6060  
Db 6050 ACAGGACTTCACTGAGGATGGCACCTCTTCAACACCTTCTACCTGGGCACTGGCCGCGAG 6109  
Qy 6061 GGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAA 6120  
Db 6110 GGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAA 6169  
Qy 6121 GGTGAGTTTCACTATGACGAGACGGCAGGCTGCTGAAGACCATCAACCTACAGAATGA 6180  
Db 6170 GGTGAGTTTCACTATGACGAGACGGCAGGCTGCTGAAGACCATCAACCTACAGAATGA 6229  
Qy 6181 GGGCTTCACTGACCATCCGCTACCGTACAGATTGGGCCCTGATTGACCGACAGATCTT 6240  
Db 6230 GGGCTTCACTGACCATCCGCTACCGTACAGATTGGGCCCTGATTGACCGACAGATCTT 6289

QY	6241	CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGGTTTTTGACTACAACACTATGACAAACAGCTT	6300
DB	6290	CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGGTTTTTGACTACAACACTATGACAAACAGCTT	6349
QY	6301	CCGGGTGACAGCATGCAGGCTGTGATCAACAGAGACCCCACTGCCCAATTGATCTCTATCG	6360
DB	6350	CCGGGTGACAGCATGCAGGCTGTGATCAACAGAGACCCCACTGCCCAATTGATCTCTATCG	6409
QY	6361	CTATGATGATGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGA	6420
DB	6410	CTATGATGATGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGA	6469
QY	6421	CATTAAACCAGATCATCACACAGCTGTGATGACCCACACCAAGCAATTTTGATGTCATATGG	6480
DB	6470	CATTAAACCAGATCATCACACAGCTGTGATGACCCACACCAAGCAATTTTGATGTCATATGG	6529
QY	6481	CAGGATGAAGGAAGTGCAGTATGAGATCTTCGGCTCGCTCATGTACTGTGATGACCGTCCA	6540
DB	6530	CAGGATGAAGGAAGTGCAGTATGAGATCTTCGGCTCGCTCATGTACTGTGATGACCGTCCA	6589
QY	6541	GTATGATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGTAGGACCCCTACGCCAATAC	6600
DB	6590	GTATGATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGTAGGACCCCTACGCCAATAC	6649
QY	6601	CACTCGTACTCTCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGA	6660
DB	6650	CACTCGTACTCTCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGA	6709
QY	6661	CAAGCCACTCTGGCGCTACAGCTACGACTACGGTATGACATCCGCAACCGCATCACTCGGCTGG	6720
DB	6710	CAAGCCACTCTGGCGCTACAGCTACGACTACGGTATGACATCCGCAACCGCATCACTCGGCTGG	6769
QY	6721	GAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCAACCGCATCACTCGGCTGG	6780
DB	6770	GAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCAACCGCATCACTCGGCTGG	6829
QY	6781	TGACGTGCAATACAAAGATGGATGAGGATGGGCTTCCTGAGGCAGCGGGCGGTGATATCTT	6840
DB	6830	TGACGTGCAATACAAAGATGGATGAGGATGGGCTTCCTGAGGCAGCGGGCGGTGATATCTT	6889
QY	6841	TGAGTACAACTCAGCTGGCCTGCTCATCAAGGCTACAACCGGGTGGCAGCTGGAGTGT	6900
DB	6890	TGAGTACAACTCAGCTGGCCTGCTCATCAAGGCTACAACCGGGTGGCAGCTGGAGTGT	6949
QY	6901	CAGGTACCGCTACGATGGCCTGGGGCGCGGTGTCCAGCAAGAGCAGCACAGCCACCA	6960
DB	6950	CAGGTACCGCTACGATGGCCTGGGGCGCGGTGTCCAGCAAGAGCAGCACAGCCACCA	7009
QY	6961	CCTGCAGTTCTTCTATGTCAGACCTGACCAACCCCAAGGTCAACCCCTGTACAACCA	7020
DB	7010	CCTGCAGTTCTTCTATGTCAGACCTGACCAACCCCAAGGTCAACCCCTGTACAACCA	7069
QY	7021	CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGA	7080
DB	7070	CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGA	7129
QY	7081	GCTGAGCAGTGGTATGAGTTTTTACATAGCTTTGTGACAAACATCGGGACCCCTCTTGCTGT	7140
DB	7130	GCTGAGCAGTGGTATGAGTTTTTACATAGCTTTGTGACAAACATCGGGACCCCTCTTGCTGT	7189
QY	7141	CTTTTAGTGGACACGGTTTGATGATCAAGCAAAATCCTGTACACAGCCTATGGGAGATCTA	7200
DB	7190	CTTTTAGTGGACACGGTTTGATGATCAAGCAAAATCCTGTACACAGCCTATGGGAGATCTA	7249
QY	7201	CATGGATACCAACCCCAACTTTTCAGATCATATAGGCTACCATGGTGGCCTCTATGATCC	7260
DB	7250	CATGGATACCAACCCCAACTTTTCAGATCATATAGGCTACCATGGTGGCCTCTATGATCC	7309
QY	7261	ACTCACCAAGCTTGTCCACATGGGCCGGGAGATTATGATGTGCTGGCCCGGACGCTGGAC	7320
DB	7310	ACTCACCAAGCTTGTCCACATGGGCCGGGAGATTATGATGTGCTGGCCCGGACGCTGGAC	7369

QY	7321	TAGCCACAGACCAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCCTTTAATCT	7380
Db	7370	TAGCCACAGACCAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCCTTTAATCT	7429
QY	7381	CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCAGGACATCAAGTGCCTTCATGAC	7440
Db	7430	CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCAGGACATCAAGTGCCTTCATGAC	7489
QY	7441	AGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTCATCCCTGGTTA	7500
Db	7490	AGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTCATCCCTGGTTA	7549
QY	7501	TCCCAAAACCAGACATGGATGCCATGGAAACCTCCTACGAGCTCATCCACACACAGATGAA	7560
Db	7550	TCCCAAAACCAGACATGGATGCCATGGAAACCTCCTACGAGCTCATCCACACACAGATGAA	7609
QY	7561	AACGACAGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA	7620
Db	7610	AACGACAGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA	7669
QY	7621	GCTCAAGCCCTTTGTCACTTAAACGGTTTGACCAAGTCTATGGCTCCACAATCACCAG	7680
Db	7670	GCTCAAGCCCTTTGTCACTTAAACGGTTTGACCAAGTCTATGGCTCCACAATCACCAG	7729
QY	7681	CTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGCACTCCAGCGGCTCAGTCTTTGGCAAGGG	7740
Db	7730	CTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGCACTCCAGCGGCTCAGTCTTTGGCAAGGG	7789
QY	7741	GGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAACAGATCATCAGTGTGGCCAATGA	7800
Db	7790	GGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAACAGATCATCAGTGTGGCCAATGA	7849
QY	7801	GGATGGCGAAGGTTGCTGCCATCTTGAACCAAGCCCACTACCTAGAGAACCTGCACTT	7860
Db	7850	GGATGGCGAAGGTTGCTGCCATCTTGAACCAAGCCCACTACCTAGAGAACCTGCACTT	7909
QY	7861	CACCAATTGATGGGTGGATACCCATTACTTTGTGAACCAGGACCTTCAGAAAGGTGACCT	7920
Db	7910	CACCAATTGATGGGTGGATACCCATTACTTTGTGAACCAGGACCTTCAGAAAGGTGACCT	7969
QY	7921	GGCCATCTGGCCCTCAGTGGGGCGGCGAACCTGGAGAATGGGTCAACGTCACCTGT	7980
Db	7970	GGCCATCTGGCCCTCAGTGGGGCGGCGAACCTGGAGAATGGGTCAACGTCACCTGT	8029
QY	7981	GTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA	8040
Db	8030	GTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA	8089
QY	8041	GTACGGGGACTGTGCTTGAACACACAGCTACGGGACAAACGTTGGATGAGGAGAAGGCACG	8100
Db	8090	GTACGGGGACTGTGCTTGAACACACAGCTACGGGACAAACGTTGGATGAGGAGAAGGCACG	8149
QY	8101	GGTCCTGGAGCTGGCCCGGCAGAGAGCCGTGCGCCAAAGCGTGGGCCCGCGAGCAGCAGAG	8160
Db	8150	GGTCCTGGAGCTGGCCCGGCAGAGAGCCGTGCGCCAAAGCGTGGGCCCGCGAGCAGCAGAG	8209
QY	8161	ACTGGGGGAAGGGAGGAAGGCCCTGCGGGCCTGGACAGAGGGGGAGAAAGCAGAGGTGCT	8220
Db	8210	ACTGGGGGAAGGGAGGAAGGCCCTGCGGGCCTGGACAGAGGGGGAGAAAGCAGAGGTGCT	8269
QY	8221	GAGCACAGGCGGTTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCC	8280
Db	8270	GAGCACAGGCGGTTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCC	8329
QY	8281	AGAACTGTACACAGCGCCCAACAACATCCACTTCATGAGACACAGCGAGATGGGCCGGAG	8340
Db	8330	AGAACTGTACACAGCGCCCAACAACATCCACTTCATGAGACACAGCGAGATGGGCCGGAG	8389
QY	8341	GTGACAGAGAGGAC	8354
Db	8390	GTGACAGAGAGGAC	8403





Db 3832 TTTACACATAGTCACAGTCCAGCACACAAATACTACTGGCCACAGACCCCATGAGTGG 3891  
QY 3904 GGCGTCTTCTCTTCTGACAGCAACAGCCGGGGTCTTTAAATCAAGTCACTGTGGT 3963  
Db 3892 GGCGTCTTCTCTTCTGACAGCAACAGCCGGGGTCTTTAAATCAAGTCACTGTGGT 3951  
QY 3964 GGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAAGTCC 4023  
Db 3952 GGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAAGTCC 4011  
QY 4024 CTTTGATGACACTCGCTGCGGGATGGTGGAAAGCCACAGAAAGCCACACTCACCAATCC 4083  
Db 4012 CTTTGATGACACTCGCTGCGGGATGGTGGAAAGCCACAGAAAGCCACACTCACCAATCC 4071  
QY 4084 CAGGGT-----ATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCAC 4134  
Db 4072 CAGGGTCCCCCAGGCATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCAC 4131  
QY 4135 CATGATCAGACGATCGATCAGAAATGGGATCATCTCCACCTGCTCGGCTCTAATGATCT 4194  
Db 4132 CATGATCAGACGATCGATCAGAAATGGGATCATCTCCACCTGCTCGGCTCTAATGATCT 4191  
QY 4195 CACATCAGCCCGCCACTCAGCTGTGATTCTGTCTATGGATATTTCACAGGTAAGA---- 4249  
Db 4192 CACATCAGCCCGCCACTCAGCTGTGATTCTGTCTATGGATATTTCACAGGTAAGACAGGT 4251  
QY 4250 ----CTGGAGTGGCCACACAGACTTAGCCATCAACCCAAATGGACAACTCACTTATGTCCT 4305  
Db 4252 TCACCTGGAGTGGCCACACAGACTTAGCCATCAACCCAAATGGACAACTCACTTATGTCCT 4311  
QY 4306 CGACAACAATGTGTCCTGCAAAATCTCTGAAAACACACAGGTGGCAATTGTCGCCGGGAG 4365  
Db 4312 CGACAACAATGTGTCCTGCAAAATCTCTGAAAACACACAGGTGGCAATTGTCGCCGGGAG 4371  
QY 4366 GCCATGCACTGCCAGGTCCCTGGCATTGACCACTTCTGTAAAGCAAGTGGCCATCCA 4425  
Db 4372 GCCATGCACTGCCAGGTCCCTGGCATTGACCACTTCTGTAAAGCAAGTGGCCATCCA 4431  
QY 4426 CGCAACCTTGAGTCAGCCACCGCTTTGGCTGTTTCAACAATGGGGTCTGTATATTC 4485  
Db 4432 CGCAACCTTGAGTCAGCCACCGCTTTGGCTGTTTCAACAATGGGGTCTGTATATTC 4491  
QY 4486 TGAGACTGATGAGAAAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTC 4545  
Db 4492 TGAGACTGATGAGAAAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTC 4551  
QY 4546 ACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTFAAAAATGATGCCAACTGTGTTT 4605  
Db 4552 ACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTFAAAAATGATGCCAACTGTGTTT 4611  
QY 4606 TTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTAAATACCCCATCTTCTTGGCTGT 4665  
Db 4612 TTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTAAATACCCCATCTTCTTGGCTGT 4671  
QY 4666 GTGTGCTGATGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTATCCG 4725  
Db 4672 GTGTGCTGATGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTATCCG 4731  
QY 4726 GAAGAACAAGCCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCA 4785  
Db 4732 GAAGAACAAGCCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCA 4791  
QY 4786 GGAGCTCTATCTGTTTATACCCACCCAGAACCTGTACACCCAAAGCCTGCCACAGG 4845  
Db 4792 GGAGCTCTATCTGTTTATACCCACCCAGAACCTGTACACCCAAAGCCTGCCACAGG 4851  
QY 4846 AGACTACCTGTACAACTTCACTACCTGGGAGCGGACATCACTCATCACAGACAA 4905  
Db 4852 AGACTACCTGTACAACTTCACTACCTGGGAGCGGACATCACTCATCACAGACAA 4911  
QY 4906 CAATGGCAACATGGTAAATGTCCGCCGAGACTTACTGGGATGCCCTCTGGCTGGTGGT 4965  
Db 4912 CAATGGCAACATGGTAAATGTCCGCCGAGACTTACTGGGATGCCCTCTGGCTGGTGGT 4971

QY 4966 CCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGAC 5025  
Db 4972 CCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGAC 5031  
QY 5026 CACACAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAAC 5085  
Db 5032 CACACAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAAC 5091  
QY 5086 CAAAAGCAATGAAAACGGATGGACAACATTTTATGAGTACGACAGCTTTTGGCCGCTGAC 5145  
Db 5092 CAAAAGCAATGAAAACGGATGGACAACATTTTATGAGTACGACAGCTTTTGGCCGCTGAC 5151  
QY 5146 AAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCACT 5205  
Db 5152 AAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCACT 5211  
QY 5206 GCATGTCCAGGTAGAGACCTCCAGCAAGGATGATGTCAACATAACCAACCTGTCTGC 5265  
Db 5212 GCATGTCCAGGTAGAGACCTCCAGCAAGGATGATGTCAACATAACCAACCTGTCTGC 5271  
QY 5266 CTAGGCGCCTTCTACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGC 5325  
Db 5272 CTAGGCGCCTTCTACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGC 5331  
QY 5326 CGATGGCTCCTTTCGGCTGCTGTCGCAACCGCATGGAGTGGCGTGCAGACTGAGCC 5385  
Db 5332 CGATGGCTCCTTTCGGCTGCTGTCGCAACCGCATGGAGTGGCGTGCAGACTGAGCC 5391  
QY 5386 CCACTTGTGGTGGCAACCGTCAACCCACCGTGGGCAAGAGGAATGTCAACCTGCCCAT 5445  
Db 5392 CCACTTGTGGTGGCAACCGTCAACCCACCGTGGGCAAGAGGAATGTCAACCTGCCCAT 5451  
QY 5446 CGACAAACGGCCTCAACCTGGTGGAGTGGCGCCACGCAAGAGCAGGCTCGGGCCAGGT 5505  
Db 5452 CGACAAACGGCCTCAACCTGGTGGAGTGGCGCCACGCAAGAGCAGGCTCGGGCCAGGT 5511  
QY 5506 CACTGTCTTTGGGCGCCGCTGCGGGTG-----CACAAACCGAATCTCTATCTCT 5556  
Db 5512 CACTGTCTTTGGGCGCCGCTGCGGGTGCTCCAGGTTCAACACCGAATCTCTATCTCT 5571  
QY 5557 GGACTTTGATCGCGTAAACACGCAAGAGAAATCTATGATGACCAACCGCAAGTTTCAACCT 5616  
Db 5572 GGACTTTGATCGCGTAAACACGCAAGAGAAATCTATGATGACCAACCGCAAGTTTCAACCT 5631  
QY 5617 TCGGATTTCTGTACGACAGCGGGGGCGCCAGCCTCTGTTCAACAGCAGCAGGCTGAA 5676  
Db 5632 TCGGATTTCTGTACGACAGCGGGGGCGCCAGCCTCTGTTCAACAGCAGCAGGCTGAA 5691  
QY 5677 TGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCAT 5736  
Db 5692 TGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCAT 5751  
QY 5737 GTCTGAAAGAAATGGAATACGACAGCGGGCGGCATCACATCCAGGATCTTCTGTGATGG 5796  
Db 5752 GTCTGAAAGAAATGGAATACGACAGCGGGCGGCATCACATCCAGGATCTTCTGTGATGG 5811  
QY 5797 GAAAGACATGGAGCTACACATACTTAGAGA-----AGTCCATGGTGTGCTACT 5844  
Db 5812 GAAAGACATGGAGCTACACATACTTAGAGAAGGAGGAGGTGTCCAGTCCATGGTGTGCTACT 5871  
QY 5845 ACACAGCCAGAGGAGTATATCTTTGAGTTCGACAAGAAATGACCGCCTCTCTTCTGTGAC 5904  
Db 5872 ACACAGCCAGAGGAGTATATCTTTGAGTTCGACAAGAAATGACCGCCTCTCTTCTGTGAC 5931  
QY 5905 GATGCCCAACGTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAA 5964  
Db 5932 GATGCCCAACGTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAA 5991  
QY 5965 CATCTATCAGCCCTGAGGGCAATGCCTCAGTCAACAGGACTTCACTGAGGATGGGCA 6024  
Db 5992 CATCTATCAGCCCTGAGGGCAATGCCTCAGTCAACAGGACTTCACTGAGGATGGGCA 6051

Qy	6025	CCTCCTTACACCTTTCTACCTGGGCACTGGCCGACAGGTGATATACAAGTATGGCAAACT	6084
Db	6052		6111
Qy	6085	GTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAGTTTCACCTATGACGAGAC	6144
Db	6112		6171
Qy	6145	GGCAGGCATGCTGAAGACCAATCAACCTACAGAAATGAGGGCTTCACCTGCACCATCCGCTA	6204
Db	6172		6231
Qy	6205	CCGTCAGATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGT	6264
Db	6232		6291
Qy	6265	CAACGCCCGTTTGTACTACAACATATGACAAACAGCTTCCGGGTGACCAAGCATGCAGGCTGT	6324
Db	6292		6355
Qy	6325	GATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTCAAGCAAGAC	6384
Db	6352		6411
Qy	6385	AGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATATGACATTAACCAAGATCATCACACAGC	6444
Db	6412		6471
Qy	6445	TGTCATGACCCACACCAAGCATTTTGTATGTCATATGGCAGGATGAAGGAAGTGCAGTATGA	6504
Db	6472		6531
Qy	6505	GATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGT	6564
Db	6532		6591
Qy	6565	GAAGAAGGAGCTGAAGGTAGGACCTACGCCAATACCCTCGCTACTCTCTATGAGTATGA	6624
Db	6592		6651
Qy	6625	TGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCAATCTGGCGCTACAGCTA	6684
Db	6652		6711
Qy	6685	CGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCAACGGTCAACCACT	6744
Db	6712		6771
Qy	6745	ACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGA	6804
Db	6772		6831
Qy	6805	GGATGGCTTCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGTGGCCCTGCT	6864
Db	6832		6891
Qy	6865	CATCAAGGCCTACAAACCGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCCTGGG	6924
Db	6892		6951
Qy	6925	GCGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCCTGCAGTTCTTCTATGCAGACCT	6984
Db	6952		7011
Qy	6985	GACCAACCCACCAAGGTCAACCCACCTGTACAACCACTCCAGCTCTGAGATCACCTCCCT	7044
Db	7012		7071
Qy	7045	CTACTACGACTTGGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGTAGTGGTTTTA	7104
Db	7072		7131
Qy	7105	CATAGCTTGTGACAAACATCGGGACCCCTCTTGTCTGTCTTTAGTGGAAACAGGTTTGATGAT	7164

7132	CATAGCTTGTGACAA	CATCGGGACCCCTCTTGCTGTCTTTAGTGGAACAGGTTTGATGAT	7191
7165	CAAGCAAATCCTGTACACAGCCCTATGGGAGATCTACATGGATACCAACCCCAACTTTCA	7224	
7192	CAAGCAAATCCTGTACACAGCCCTATGGGAGATCTACATGGATACCAACCCCAACTTTCA	7251	
7225	GATCATATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGG	7284	
7252	GATCATATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGG	7311	
7285	CCGGCGAGATTATGATGTGCTGGCCGACGCTGGACTAGCCCGAGACCACGAGCTGTGGAA	7344	
7312	CCGGCGAGATTATGATGTGCTGGCCGACGCTGGACTAGCCCGAGACCACGAGCTGTGGAA	7371	
7345	GCACCTTAGTAGCAGCAACGTCATGCCCTTTTAATCTCTATATGTTCAAAAACAACAACCC	7404	
7372	GCACCTTAGTAGCAGCAACGTCATGCCCTTTTAATCTCTATATGTTCAAAAACAACAACCC	7431	
7405	CATCAGCAACTCCAGGACATCAAGTGCCTTCATGACAGATGTTAACAGCTGGCTGCTCAC	7464	
7432	CATCAGCAACTCCAGGACATCAAGTGCCTTCATGACAGATGTTAACAGCTGGCTGCTCAC	7491	
7465	CTTTGGAATCCAGCTACAAACGTGATCCCTGGTTATCCCAAAACCAGACATGGATGCCAT	7524	
7492	CTTTGGAATCCAGCTACAAACGTGATCCCTGGTTATCCCAAAACCAGACATGGATGCCAT	7551	
7525	GGAACTCTCTACGAGCTCATCCACACACAGATGAAAAACGAGGATGGGACAAACAGCA -	7583	
7552	GGAACTCTCTACGAGCTCATCCACACACAGATGAAAAACGAGGATGGGACAAACAGCAA	7611	
7584	-----AGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA	7620	
7612	GGTAATTCCTGCACAAAGGTGCCAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA	7671	
7621	GCTCAAGGCCTTTGTACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAG	7680	
7672	GCTCAAGGCCTTTGTACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAG	7731	
7681	CTGCCAGCAGGCTCCAAAGACCAAGAAGTTTGCAATCCAGCGGCTCAGTCTTTGGCAAGGG	7740	
7732	CTGCCAGCAGGCTCCAAAGACCAAGAAGTTTGCAATCCAGCGGCTCAGTCTTTGGCAAGGG	7791	
7741	GGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTGGCCAATGA	7800	
7792	GGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTGGCCAATGA	7851	
7801	GGATGGCGAAGGTTGCTGCCATCTTGAACCAATGCCACTACCTAGAGAACCTGCACCT	7860	
7852	GGATGGCGAAGGTTGCTGCCATCTTGAACCAATGCCACTACCTAGAGAACCTGCACCT	7911	
7861	CACCATTTGATGGGTGGATACCCATTACTTTGTGAACACAGGACCTTCAGAAGGTGACCT	7920	
7912	CACCATTTGATGGGTGGATACCCATTACTTTGTGAACACAGGACCTTCAGAAGGTGACCT	7971	
7921	GGCCATCCTGGSCCTCAGTGGGGGGCGGGAACCTTGAGAAATGGGGTCAACGTCACTGT	7980	
7972	GGCCATCCTGGSCCTCAGTGGGGGGCGGGAACCTTGAGAAATGGGGTCAACGTCACTGT	8031	
7981	GTCCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA	8040	
8032	GTCCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA	8091	
8041	GTAACGGGCACTGTGCTTTGAACACACCGCTACGGGACAAAGTTGGATGAGGAGAGGCACG	8100	
8092	GTAACGGGCACTGTGCTTTGAACACACCGCTACGGGACAAAGTTGGATGAGGAGAGGCACG	8151	
8101	GGTCTTGAGCTGGCCCCGGCAGAGAGCCGTGGCCAAAGCGTGGGCCCCGGAGCAGCAGAG	8160	
8152	GGTCTTGAGCTGGCCCCGGCAGAGAGCCGTGGCCAAAGCGTGGGCCCCGGAGCAGCAGAG	8211	
8161	ACTGCGGGAAGGGAGGAAGGCCCTGCGGGCCTTGGAACAGAGGGGGAAGCAGCAGGTGCT	8220	

Db 8212 ACTGCGGGAAGGGAGGAGGCGCTCGGGCCCTGGACAGAGGGGAGAGCAGCAGGTGCT 8271

Qy 8221 GAGCACAGGCGGGTGTCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8280

Db 8272 GAGCACAGGCGGGTGTCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8331

Qy 8281 AGAACTGTCAAGCAGCGCCCAACAACATCCACTTCATGACAGAGCAGAGATGGCGCGGAG 8340

Db 8332 AGAACTGTCAAGCAGCGCCCAACAACATCCACTTCATGACAGAGCAGAGATGGCGCGGAG 8391

Qy 8341 GTGACAGAGGAC 8354

Db 8392 GTGACAGAGGAC 8405

RESULT 4

AB025413

LOCUS

DEFINITION Mus musculus mRNA for Ten-m4, complete cds.

ACCESSION AB025413

VERSION AB025413.1 GI:4760781

KEYWORDS Ten-m4.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (sites)

Ohashi, T., Zhou, X., Feng, K., Richter, B., Moergelin, M., Perez, M.T., Su, W., Chiquet-Ehrismann, R., Rauch, U. and Faessler, R.

Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues

J. Cell Biol. (1999) In press

Ohashi, T.

Direct Submission

Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8252, Japan

(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128, Fax: +81-86-222-7768)

FEATURES

source

1. .8585

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="Balb/c"

/db\_xref="taxon:10090"

/tissue\_type="brain"

/dev\_stage="adult"

1. .8585

/gene="ten-m4"

188. .8503

/gene="ten-m4"

/codon\_start=1

/product="Ten-m4"

/protein\_id="BAA77399.1"

/db\_xref="GI:4760782"

/translation="MDVKERKPYRLTRRRDAERRYTSADSEEGKGPQKYSVSSET LKAYDQARLAYSRVKDMVPQEAEEFCRTGTNFTLRELIGEMTPPHGLTYRDTIGL PHCYSGASSDADLEADTVLSPHPVRLWGRSTRSGRSCLSRANSNLTLTTEHE NTEHDHPSLQNHPLRLTPPPPLPHATPNQHAASINSLNRGNFTPRSNPSAPT DH SLGEPAPGSAQEPHTAQDNWLLNSNIPLETNLGKQPFGLTQDNLIEMDILSARH DSGYSDGHFLFKPGTSPLEFCTTSPGYPLTSSTVYSPPPPLPRSTFSRPAFNLKPS KYCNWKAALSAILISATLVLLAYFVAMHLFGLNWHLPQMEGQMOMYEITEDTASSW PVPTDVSLYPSGGTGLETPDRKGAAEGKPSLFPEDSFIDSGEIDVGRRASQI PP GTFWRSQVFIHPVHLKFNVLGKAALVGIYGRKGLPPSHTFQDFVELLDGRRLLTQE ARSLGPQRQSRGPVPSSSHETGFIQYLDLSGIWHLAFYNDGKSEVSVFLTTAESVD NCPNSCYNGDNCISGTHCFLGPDGCRASCPVLCSNGQYMKGRCLCHSGWKGA E CDVPTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDCMDPTCSSRGVCVRGECH SVGWGNTCETPRATCLDQCSHGTFLPDTGLCNCDPSWTGHDCSI EICAADCGHGV CVGGTCRCDGWMGAACDQACHPRCAEHGTCRDGKCESPGWNGEHTIAHLDV RV KEGCPGLCNGNRCCTLDLNGWHVCVQLGWRGTCDTSMETCGCDKNDGDGLVDCMD PDCCQLPLCHVNPCLGSPDPLDI IETQAPVSQNLNPFYDRIKFLVGRDSTHSPG ENPFDGGHACVIRGQVMTSDGTPLVGVNLSFINNPLFGYITISRDGFSFDLVINGGISI

ILRFERAPFITQEHTLWLPWDRFFVMEIIVMRHEENEIPSCDLSNFARPNPVWSPSPPL

TSFASSCAEKGPVPEIQALQEEIIVIAQKMRSLYSLSRTPGYKSLRISLTHPTIPF

NLMKVHLMVAVEGLFRKWFAPDLASYFTWDTVYNQKVFGEFAFVSVGEYES

CPDLILWEKRTAVLQGEIDASKLGLWSLDKHALNIIQSGILHKGNGENQFVSQPPV

IGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFNIRRIFFPSGNV

TNILEMRNKDFRSHSPAHKYLYLATDPMGAVFLSDTNSRRVFKVSTTVTKDLVQNS

EVVAGTQDQCLPFDDTRCGDGGKATEATLTNPRGTLVDKFLIYFVDTGTMIRVDQNG

IISTLLGSDNLTARPLSCDSVMEISQVLEWPTDLAINPMDNSLYVLDDNNVVLQISE

NHQRIVAGRPMHCQVPGIDHFLLSKVHATLESATALAVSHNGVLYIAETDEKKN

RIRQVTTSGEISLVAGAPSGCDKNDANDCFSGDDGYAKADAKLNTPPSSLAVCADGEL

YVADLGNIRIRIRKKNKPFNTQNMVELSSPIDQELYLFDTSKHLYTQSLPTGDYLY

NFTYTGDGDIITHITDNNGMVNRDSTGMPLWLVPDQVYVWVTMTNSALRSVTTQ

GHELAMTYHNGSGLLATKSNENGWTFYEDSFGRLTNTVTFPTQVSSFRSDTSSV

HVQVETSSKDDVTITNLSASGAFYTLQDQVRNSYIIGADGSLRLLLANGMEVALQT

EPHLLAGTVNPTVGRKRVNLTPIDNLGLNLEWRQRKEQGVTVFGRRLRVHNRNLLS

LDLDFRVTRTEKIYDDHRKFTLRILYDQAGPSLWSLSPSSRLNGVNVTVSPGGHIAGIQ

GIMSERMEYDQAGRITSRIFADGKMSYTYLEKSMVLHLSHQYIFEFKDNDRSSV

TMPNVARQTLTETIRSVGYRNIYQPPENASVIQDFTEDGHLHLYLGTGRRVIYKY

EGLSKLAETLYDTTKVSFTYDETAGMLKTNLQNEGFTTIRYQIGPLIDRQIFRFT

EGGMVNARFDVNDNSFRVTSMAVINETPLPIDLYRDDVSGKTEQFGKFGVIYDI

NQIITAVMTHTKHFDAYGRMEVQYEIFRSLMYTMVQYDNMVRVVKELKVGPGYAN

TTRYSEYDADQLOQTVSINDKPLWRSYDLNGLNHLHLLSPGNSARLTPLRYDLDRIT

RLGDVQYKMEDEGFLRQDGDVFEYNAGLLIKAYNRASGWSVRYRYDGLGRRVSSKS

SHSHLQFFYADLTNPTKVTHLYNHSSEITSLYLDQLQHLFAMELSSGDEFYIACDN

IGTPLAVFSGTGLMIKQILYTAGYEIYMDTNPNFQIIIGYHGLYDPLTKLVHMRDRD

YDLAGRWTSPDHMLKRLSSNSIVPHLYMFKNNPISNSDIKCFMTDVNSWLLTF

GFQHNVI PGYKPDTDAMEPSYELVHTQMTQEWDSKSLGVQCEVQKQKAFVTL

ERFDQLYGSTITSCQAPETKKFASGSI FGKGVKFAKLDGRVTTDIISVANEDGRII

AAILNNAHLYENLHFTIDGVDTHYFKPSPSEGDLAILGLSGRRRTLENGVNVTVSQI

NTMLSGRTRRYTDIQLQYRALCLNTRYCTTVDEKRVLELARQRAVRQAWAREQRL

REGEEGLRAWTDGEKQQLVNTGRVQYDGFVTSVEQYPELSDSANNIHFMRQSEMR

R"

ORIGIN

Query Match 83.5%; Score 6975.4; DB 10; Length 8585;

Best Local Similarity 89.7%; Pred. No. 0;

Matches 7502; Conservative 0; Mismatches 851; Indels 6; Gaps 1;

Qy 2 TTTGTGGATGTGGAGGAGCGCGCGGAGCGCGCGCTACACCGCTCGTCCGCGGACAGC 61

Db 155 TTCGTGGATGTGGAAGAGCGAGAGCCAGAGGCCATGGACCTGGAAGGAGGAAGCCCTAT 214

Qy 62 CGCTCGCTGACCCCGCGCGGAGCGCGCGCTACACCGCTCGTCCGCGGACAGC 121

Db 215 CGCTCGCTGACCCCGCGCTCGCATGTCTGAGCGCGCTACACCGCTCATCGGACAGAGC 274

Qy 122 GAGGAGGGCAAGCCCGCGAGAAATCGTACAGCTCCAGCGAGACCCCTGAAGGCCTACGAC 181

Db 275 GAGGAGGGCAAGGGCCACAGAAATCCTACAGCTCCAGTGAGACCCCTGAAGGCATATGAC 334

Qy 182 CAGGACGCCCGCTAGCCTATGGCAGCGCGCTCAAGGACATTTGTCCGCGAGGCGCGAG 241

Db 335 CAAGATGCTCGCCCTAGCCTACGGCAGCGCGCTCAAGGACATGGTACCACAGGAGCGCGAG 394

Qy 242 GAATTCTCGCGCACAGGTGCCAACTTCACTCCCTCGCGGAGCTGGGGCTGGAAGAAAGTAACG 301

Db 395 GAGTTCTCGCGCACAGGCACATAATTTCACTCTTCTGTAAGTGGGACTGGGAGAGATGACG 454

Qy 302 CCCCCCTACGGGACCCCTGTACCGGACAGACATTTGGCCTGCCCCAAATGGGCTACTCCATG 361

Db 455 CCCCCCTATGGGACTCTCTACAGGACAGACATCGGCGCTCCACACTGTGGCTATTCCATG 514

Qy 362 GGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTGTCTCCCTGAGCACCCCGTG 421

Db 515 GGGGCCAGCTCTGATGCAGACTTGAAGCAGACACTGTGTCTATCCCTGAAACACCCAGTG 574

Qy 422 CGTCTGTGGGCGCGGACACACGGTCAAGGCGCGAGCTCCTGCCTGTCCAGCCGGCCCAAT 481

Db 575 CGCCTGTGGGCGCGGACACACGGTCAAGGCGCGAGTTCTCTGCCTGTCCAGCCGGCCCAAC 634

Qy 482 TCCAATCTCACACTCACCGGACACCGGACATGAAACACTGAGACTGATCATCCGGGCGGC 541

Db 635 TCCAACCTCACACTCACGGACACAGAGCATGAGAACACAGAACCCGATCATCCAGCAGC 694







||||| 7235 ATGGAGCTGAGCAGTGGCGATGAGTTTACATAGCTTGTGACAAACATCGGGACCCCGCTT 7294  
||||| 7136 GCTGTCTTTAGTGAACAGGTTTGTATGATCAAGCAAAATCCTGTACACAGCCTATGGGAG 7195  
||||| 7295 GCTGTCTTCAAGTGAACACTGGCTTGATGATTAAGCAGATCCTATACACGGCCTATGGGAG 7354  
||||| 7196 ATCTACATGGATACCAACCCCACTTTAGATCATATAGCTACCATGGTGGCCTCTAT 7255  
||||| 7355 ATCTACATGGACACAAACCCCACTTCCAGATCATATCGGCTACCAACGGCGGCTCTAT 7414  
||||| 7256 GATCCACTCACCAAGCTTGTCCACATGGCGCGGAGATTATGATGTGCTGGCCGACGC 7315  
||||| 7415 GATCCACTCACCAAGCTTGTCCACATGGCGCGGAGATTATGATGTGCTGGCTGACGC 7474  
||||| 7316 TGGACTAGCCAGACACGAGCTGTGGAAGCACCTTTAGTAGCAGCAACGTCATGCTTTT 7375  
||||| 7475 TGGACAAAGCCAGACCATGAATCTGGAACCGCTGTAGTAGCAACAGCATCGTGCCTTT 7534  
||||| 7376 AATCTCTATATGTTCAAAAACCAACCCCATCAGCAACTCCCAGGACATCAAGTCTTC 7435  
||||| 7535 CATCTCTACATGTTTAAGAACACAAACCCCATCAGCAACTCTCAGGACATCAAGTCTTC 7594  
||||| 7436 ATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACAAACGTCATCCCT 7495  
||||| 7595 ATGACAGATGTCAACAGCTGGCTCCTCACCTTTGGATTCCAGCTGCACAAAGTATACCT 7654  
||||| 7496 GGTATCCCAACACAGACATGATGATGCCATGGAACCCCTCCTACGAGCTCATCACACACAG 7555  
||||| 7655 GGTATCCCAAGCCAGACACAGATGCCATGGAACCCCTCCTACGAGCTCGTACACACAG 7714  
||||| 7556 ATGAAAACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAG 7615  
||||| 7715 ATGAAAACGCAGGATGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTTCAG 7774  
||||| 7616 AAGAGCTCAAGGCTTTGTACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATC 7675  
||||| 7775 AAGCAACTCAAGGCTTTGCTTACCTTAGAACGGTTTGACAGCTCTACGGCTCGACCATC 7834  
||||| 7676 ACCAGCTGCCAGAGGCTCCAAAGACCAAGTGTGCTATCCAGCGGCTCAGTCTTTGGC 7735  
||||| 7835 ACCAGCTGCCAACAGGCTCCAGGCTGAGACAAAGTGTGCTCCAGTGTTCATCTTTGGC 7894  
||||| 7736 AAGGGGTCAAGTTTGCCTTGAAGGATGGCGAGTGACCAACAGACATCATCAGTGTGGCC 7795  
||||| 7895 AAGGGGTCAAGTTTGCCTTGAAGGATGGTCGAGTGACCACTGACATCATCAGTGTGGCC 7954  
||||| 7796 AATGAGGATGGCGAAGGTTGTGCCATCTTGAACCAATGCCACTACCTAGAGAACCTG 7855  
||||| 7955 AATGAGGATGGCGGAGGATGCGAGCCATCTTGAACAATGCTCACTATCTAGAGAACCTT 8014  
||||| 7856 CACTTCACCATTTGATGGGTGGATACCCATTACTTTGTGAACCAAGGACCTTCAGAAGGT 7915  
||||| 8015 CACTTCACCATTTGATGGGTGGACACCCACTACTTTGTGAACCAAGGACCTTCGGAAGGT 8074  
||||| 7916 GACCTGGCCATCCTGGGCTCAGTGGGGGGCGGCAACCTGGAGATGGGGTCAACGTC 7975  
||||| 8075 GACCTGGCCATCCTGGGCTCAGTGGGGGGCGGAGAACCTGGAGATGGGGTCAACGTC 8134  
||||| 7976 ACTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAG 8035  
||||| 8135 ACTGTGTCCAGATCAACACCATGCTCAGTGGCAGGACTAGACGCTACACTGACATCCAG 8194  
||||| 8036 CTCCAGTACGGGGCACTGTGCTTGAACACACAGCTACGGGCAACCTGGAGATGGGAGAG 8095  
||||| 8195 CTGCAGTACAGGGCGCTGTGCCTGAACACCCGCTACGGGACAAACAGTGGATGAGGAAAAG 8254  
||||| 8096 GCACGGGTCTTGAGCTGGCCCGGACAGAGCGCTGCGCCAAAGCGTGGGCCCCCGAGCAG 8155  
||||| 8255 GTGCGGGTGTGAGCTGGCCAGGACAGAGAGTGTGCGCCAGGCTTGGGCCCCGAGAGCAG 8314  
||||| 8156 CAGAGACTGCGGGAGAGGAGGAAGCGCTGCGGGCTTGACAGAGGGGGAGAGCAGCAG 8215  
|||||

Db 8315 CAGAGACTCGGGGAAGGGAAGAGGGCCTCGGGCCTGGACAGATGGGAAAAGCAGCAG 8374  
QY 8216 GTGCTGAGCACAGGGCGGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAG 8275  
||||| 8375 GTGCTGAACACGGGGCGGGTGCAAGGCTACGACGGCTTCTTGTGACCTCGGTGAGCAG 8434  
QY 8276 TACCCAGAACTGTGACAGACGCGCAACCAACATCCACTTCATGAGACAGACGAGATGGC 8335  
Db 8435 TACCCAGAACTGTGACAGACGCGCAACCAATATCCACTTCATGAGACAGACGAGATGGC 8494  
QY 8336 CGGAGGTGACAGAGAGGAC 8354  
Db 8495 CGAAGGTGACAGAAAGGC 8513  
  
RESULT 5  
AF059485  
LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998  
DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.  
ACCESSION AF059485  
VERSION AF059485.1 GI:3170614  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 9722)  
AUTHORS Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P.,  
Zinszner,H. and Ron,D.  
TITLE Identification of novel stress-induced genes downstream of chop  
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)  
MEDLINE 98315054  
PUBMED 9649432  
REFERENCE 2 (bases 1 to 9722)  
AUTHORS Wang,X.-Z. and Ron,D.  
TITLE Direct Submission  
JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University  
Medical Center, 550 First Ave., New York, NY 10016, USA  
FEATURES  
source  
1..9722  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NIH/Swiss"  
/db\_xref="taxon:10090"  
/cell\_line="NIH-3T3"  
/note="tunicamycin-treated"  
1..9722  
/gene="Doc4"  
/note="downstream of CHOP 4"  
583..9060  
/gene="Doc4"  
/note="similar to Drosophila melanogaster tenm/odz and human gamma-herregulin; type II transmembrane protein"  
/codon\_start=1  
/product="DOC4"  
/protein\_id="AAC31807.1"  
/db\_xref="GI:3170615"  
/translation="MDVKERKPYRSLRRRDAERRYTSADSEEGKPKQKSYSSSET  
LKAYDQARLAYSRVKDMVPQEAEEFCRTGNTFTLRELGLGEMTPPHGTLRTDGL  
PHCGYSMGASDDADLEADTVLSPHPVRLWGRSTRSRSSCLSSRANSLTLTDEHE  
NTETGAPLHCSSASSTPIEQSPSPPPANESQRRLLGNVAQPTPDSDEEEFVN  
SFLVKSASLGLVAANDHPSSLQNHPLRTPPPLPHAPTQNHHAASINSLNRNFT  
PRSNPSAPTDHLSGEPAGSAQEPHTAQNWLNSKIPVETRNLGKQPFLLGTWQDN  
LIEMDIFSARRDGYSDGHFFKPGTSPFCFTTSPGYPLTSTVYSPRPPLPRST  
FSRPAFLKPKSKYCNWKCAALSAILISATLVLLAYFVAMHLFGLNWLHPMEGQMQ  
MYEITEDTASSWPVPTDVSYPGSGTLETPDRKGKGAEGKPSLFPEDSFIDSGEI  
DVGRASQKIPPGTFRWSQVFDHPVHLKFNVSGLKALVGIYGRKGLPPSHTQLDFV  
ELLDGRRLTQEARSLGEPQSRGPVPPSSHETGFTQYLDSDGIWHLAFYNDGKESEV  
VSFLTIAESVDNCPNSCYNGDCISGTCFLGFLGDCGRASCPLVCSGNGQYMKG  
RCLCHSGWKGAECDDVPTNQCIDVACSHGTCIMGTICINPGYKGESCEVEDMDPTCS  
SRGVCVRGECHCSVWGWTNCTPRATCLDQCSHGHTFLPDTGLCNCDPSTWHDCSI  
EICAADCGGHGVCVGGTCRCEDGWMGAACDQRCAPCAEHGTCRDGKCECTPGWNGE  
HCTIEGCPGLCNGNGRCRTLDLNGWHVCVQLGWRGTGCDTSMETGCGDKNDGDLVD

CMDPCCLOPLCHVNPLCLGSPDPLDIIQETQAPVSQQNLNSFYDRIKFLVGRDSTHS  
IPGENPFDGGHACVIRGOVMTSDGTPLVGNIISFINNPLEGTYISRQDGSFDLVTNGG  
ISIIIRFERAPFITQHEITLWLPDRFFVEMETIVMRHEENETPSRDSNFRPNPVVSP  
SPLTSPASSCAEKGPVPEIQALQEEIVIAAGKMRLSYLSRTPGYKSVVRISLTGPT  
IPFNLMKVLHMAVEGRFLRWFAAAPDLSSYFIWDKTDVYNQKVFGLSEAFVSVG  
YESCPDLILWEKRTAVLQGYEIDASKLGGWSLDKXHALNIQSGILHKNGENQFVSQ  
PNVTISIMGNGRRRSISPCSCNGLADGNKLLAPVALTCGSDGSLYVGDNFYIRRIFFS  
GNTVINILEMSSHPSAKHYILATDPMGSAVFLSDTNSRRRVFKVKTIVVKDLVKNSEVVA  
GTGQOCLPDDTTRCGDGGKATEALTNPRGITVDKFLGIYFVDGTMIRRVQDNGIIST  
LLGENDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQV  
RIVAGRPMHCQVPGIDQFLLSKVAITHATLESATALAVSHNGVLYIAETDEKKINRIRQ  
VTTSGEISILVAGAPSGCDCKNDANCDFSGDDGYAKDAKLTNPSSLAVCADGELYNYV  
LGNIRIRFIRKKNPFLNTQNMYYELSSPIDQELYLEFTSGKHLTYQLSDADGELYNYV  
TGGDITHITDNNGNMNVNRDSTCMPLWLVLVDPGQVYVWVTMTNSALRSVTTQGHLE  
AMMYHNGSGLLATKSNENGWTTFEYDVSFGRLTNTVFTPTGQVSSFRSDTSSSVHVQV  
ETSSKDDVTITTHLSGSAFYTLLODQVRNSYIIGAGDGLRLLLANGMEVALQTEPHL  
LAGTVNPTVGKRNVTLPIDNGLNLVWRQKEQAGQVTVFGPPLRVHNRNLLSLDFD  
RVTRTEKIYDDHRKFTLRILIYDQAGRPSPFWSPSSRLNGVNTVSPGGHIAGIQRGIS  
ERMEYDQAGRITSIRIFADGKWSYTYLEKSMVLHLHSQROYIEFFDKNDRLSSVTMPN  
VARQLEITRSVGYRNIYQPPEGNASVIQDFTDGLLHTFLYLGTRRRVYKYGKLS  
KLAETLYDTTKVSFTYDGTAGMLKTVNLQNEGFTCTIRYQIGLIDRQIFRFTEEGM  
VNARFDYNYDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQCFKFGVYIYDINQII  
TTAVMTSHKHFDAYGRMKEVQYETFRSLMYMTVQYDNMGRVVKELKVGYPYANTTRY  
SYEYDADGQLQTVSINDKPLWRSYVDLNGNLHLLSPGNSARLTPLRYDLRLDRITRLGD  
VQYKMDGSLRLQRGDVFEYNSAGLLIKAYNRASGWSVRYRDLGTRRRVSSKSSHSH  
HLQFFYADLTNPTKVTHLYSHSSSEITSLYYDLQGLFAMELSSDGEFYIACDNIIGP  
LAVFSGELMIKQILYATGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMRDRDYDL  
AGRWTSPDHELWKRLSSNSIVPFHLYMEKNNNPISNSQDIKCFMTDVNSWLLTFFGQL  
HNYIPGPKPDDTAMEPSYELVHTQMTQEWDNKSILGVQCEVQQLKAFVTLERFD  
QLYGSTITSQQAPETKKFASSGSIFGKGVKFALKDGRVTTDIIISVANDGRRIRIAAIL  
NNAHYLNLSTCTIDGVDTHYFVKPGPSGDLAII LGLSGRRRTLENGVNVTVSQINTML  
SGRTRRYTDIQLQYRALCLNTRYGTTVDEEKVRVLELARQRAVRQAWAREQORLREG  
EGLRAWTDGEKQQVLNTRGVQYDGGFFVTSVEQYPELSDSANNIHPMRQSEMGR "

1813. .1881  
/note="putative; transmembrane-region site"  
/gene="Doc4"

misc\_feature

## ORIGIN

Query Match	79.0%;	Score 6601.2;	DB 10;	Length 9722;
Best Local Similarity	86.7%;	Pred. No. 0;		
Matches 7426;	Conservative	5;	Mismatches 874;	Indels 264; Gaps 4;
QY	2	TTTGTGGATGTGGAGAGCGCGGGCCGGAGGCCATGTGACGTGAAGGAGAGGAAGCCTTAC	61	
Db	550	TTCTGTGGATGTGGAAGACGAGAGCCAGAGGCCATGTGACGTGAAGGAGAGGAAGCCTTAT	609	
QY	62	CGCTCGCTGACCCCGCGCGCGACGCCGAGCGCCGCTACACACAGCTCGTCCGCGACAGC	121	
Db	610	CGCTCGCTGACCCCGCGCTGCGATGCTGAGCGCCGCTACACACAGCTCATCGGCAGACAGC	669	
QY	122	GAGGAGGGCAAAGCCCGCAGAAAATCGTACAGCTCCAGCGAGACCCCTGAAGGCCTACGAC	181	
Db	670	GAGGAGGGCAAAGGCCCAAGAGTCTTACAGCTCCAGTGAGACCCCTGAAGGCAATGAC	729	
QY	182	CAGGACGCCCGCCTAGCCTATGGCAGCCGCGTCAAGGACATTTGTCGCGCAGGAGGCGGAG	241	
Db	730	CAAGATGCTCGCCTAGCCTACGGCAGCCGCGTCAAGGACATGGTACCAAGGAGGCGGAG	789	
QY	242	GAATTCTGCCGCACAGGTGCCAACTTCACCCCTGCCGGGAGCTGGGCTGGAAGAAGTAACG	301	
Db	790	GAGTTCTGCCGCACAGGCACATAATTTCACTCTTCGTGTAAGTGGGACTGGGAGAGATGACG	849	
QY	302	CCCCCTCAGGGGACCCCTGTACCGGACAGACATTTGGCCCTGCCCCCAATGCGGCTACTCCATG	361	
Db	850	CCCCCTCATGGGACTCTCTACAGGACAGACATCGGCCCTCCACACACTGTGGCTATTCCATG	909	
QY	362	GGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTGCTGCTCCCTGAGCACCCCGTG	421	
Db	910	GGGGCCAGCTCTGATGCAGACCTGGAAGCAGACACTGTGCTATCCCCCTGAACACCCAGTGTG	969	
QY	422	CGTCTGTGGGGCCGGAGCAACCGGTTCAGGGCGCAGCTCTGCTGCTGTCCAGCCGGGCCAAT	481	
Db	970	CGCCTGTGGGGCCGGAGCAACCGGTTCAGGGCGCAGTTCCTGCTGCTGTCCAGCCGGGCCAAC	1029	

QY 1346 GTGGGAAGGCGAGCCTCCAGAAAGATTTCCTCTGGCACCTTTCTGGAGATCTCAAGTGTTC 1405  
Db 2110 GTGGGAGACGGGCTTCCAGAAAGATTCCGCTGGAACGTTCTGGAGATCTCAGGTGTTC 2169  
QY 1406 ATAGACCAATCCTGTGCATCTGAAATTCATATGTCTCTGGGAAAGGCGAGCCTGTGGC 1465  
Db 2170 ATAGACCACCTGTACACCTGAAGTTCAATGTGTCTCTGGGAAAGCAGCTCTGTGTCGGC 2229  
QY 1466 ATTTATGGCAGAAAAGCCCTCCCTCCTTCACATACACAGTTTCACTTTGTGGAGTGTCTG 1525  
Db 2230 ATTTATGGCAGAAAAGCCCTTCCCTCCTTCCATACTCAGTTAGACTTTGTGGAGTCTCTG 2289  
QY 1526 GATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGCCTAGAGGGACCCCGCGCCAGTCT 1585  
Db 2290 GATGGAAGAAGGCTCCTAACCCAGAGGCAAGGAGCCTGGAGGTCTCAGCGCCAATCA 2349  
QY 1586 CGGGGAATGTGCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTCAAGGA 1645  
Db 2350 CGGGGCTGTTCCTCCCATCCAGCCATGAGACCGGTTTCATCCAGTATCTGGATTCAAGGA 2409  
QY 1646 ATCTGGCACTTGGCTTTTACAAATGACGGAAAGGAGTCAGAAAGTGGTTTCTCTCACC 1705  
Db 2410 ATCTGGCACTTGGCTTTTACAAATGATGGAAGAAATGAAGTGGTTTCAATTTCTCACC 2469  
QY 1706 ACTGCCATTGAGTCGGTGGATTAAGTCCCCAGCAACTGCTATGGCAATGGTGACTGCATC 1765  
Db 2470 ACCGCCATTGAATCAGTGGATAACTGCCCCAGCAACTGCTATGGTAAACGGTGACTGCATT 2529  
QY 1766 TCTGGACCTGCCACTGCTTCTGGGTTTCTGGGCCCCGACTGTGGCAGAGCCTCCTGC 1825  
Db 2530 TCCGGACCTGCCACTGCTTCTGGGTTTCTGGGCCCCGACTGTGGCCGAGCATCCTGT 2589  
QY 1826 CCCGTGCTGTAGCGGAATGGCCAAATACATGAAGGAGATGCTTTGTGCCACAGTGGC 1885  
Db 2590 CCCGTGCTGTGAGTGGGAACCGCCAGTACATGAAGGAGAGTGCCTGTGCCACAGTGGC 2649  
QY 1886 TGGAAAGGCGTGTGAGTGGATGTGCCAACACCAACCAAGTATCGATGTGGCCTGCAGCAAC 1945  
Db 2650 TGGAAAGGCGTGTGAGTGGATGTGCCAACCAACCAAGTATCGATGTGGCCTGCAGCAGT 2709  
QY 1946 CATGGACCTGCATCAGGGACCTGCATCTGCAACCTGGCTACAAAGGGGAGAGCTGT 2005  
Db 2710 CATGGACCTGCATCAGGGACCTGCATCTGCAACCTGGCTACAAAGGGGAGAGCTGC 2769  
QY 2006 GAGGAAGTGGACTGCATGGACCCACATGTTTCAGGCGGGGTGTCTGCGTGAGAGCGGAA 2065  
Db 2770 GAGGAAGTGGACTGCATGGACCCACATGTTCCAGCAGGGGTGTTGTGTGAGAGCGGAG 2829  
QY 2066 TGCCATTGCTTTGTGGATGGGAGGCAACCACTGCGAGACCCCGAGGCAACATGCTTA 2125  
Db 2830 TGCCACTGCTGTGTGGATGGGAGGCAACCACTGCGAGACACCCAGGGCCACATGCTTG 2889  
QY 2126 GACCACTGTTTTCAGGCGCACGAACTTCTCCCGGACACCGGGCTTTGCACTGTGACCCA 2185  
Db 2890 GACCACTGCTCAGGCGCACGAACTTCTCCCGAGACACAGGACTTTGCACTGTGACCCA 2949  
QY 2186 AGCTGGACTGGACACGACTGTTTCTATCGAGATCTGTGCTGCGGACTGTGGTGGCCATGGC 2245  
Db 2950 AGCTGGACTGGACACGACTGTTTCCATTGAGATCTGTGCGGCTGACTGTGGTGGCCATGGC 3009  
QY 2246 GTGTGCGTAGGGGACCTGCCGCTGCGAGGATGGTGGATGGGGGAGCCTGCGACCCAG 2305  
Db 3010 GTGTGCGTAGGGGACCTGCCGCTGTGAGGATGGTGGATGGGGGCGCCATGCGACCAA 3069  
QY 2306 CGGGCCTGCCACCCCGCTGTGCCGAGCATGGGACCTGCCGCGACGGCAAGTGGAGTGC 2365  
Db 3070 CGGGCCTGCCACCCCGCTGTGCAGAAACCGGACCTGCCGCGACCGCAAGTGGCAATGC 3129  
QY 2366 AGCCCTGGCTGGAATGGCGAACACTGCACCATCGCTCACTATCTGGATAGGGTAGTTAA 2425  
Db 3130 ACCCCCGGCTGGAATGGAGAGCACTGCACCATC----- 3162  
QY 2426 GAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGG 2485

Db 3163 GAGGGCTGTCTGGCTTGTGCAATGGAATGGCAGATGTACCTTGGACCTGAATGGTGG 3222  
QY 2486 CACTGCGTCTGCCAGCTGGGCTGGAGAGGAGTGGCTGTGACACTTCCATGGAGACTGCC 2545  
Db 3223 CACTGTGTCTGCCAGCTGGGCTGGCGAGGAGTGGCTGTGGTGGACTGCATGGAAACGGGC 3282  
QY 2546 TCGGCTGACAGCAAAAGACAAATGATGGAGATGGCCTGTGGACTGCATGGACCTGACTGC 2605  
Db 3283 TGTGGAGACGGCAAGGACAAACGACGGAGATGGCTTGTGGACTGCATGGACCTGACTGC 3342  
QY 2606 TGCCTCCAGCCCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCCCTAACCTCTGGAC 2665  
Db 3343 TGCCTCCAGCCTCTCTGTGATGTCAACCCGCTGTGCTTAGCTCCCCCTGACCCCTGGAC 3402  
QY 2666 ATCATCCAGGAGACACAGGTCCTGTGTTCACAGCAGAACCTACACTCCTTCTATGACCGC 2725  
Db 3403 ATCATCCAGGAGACACAGGCCCCCTGTATCCAGCAGAACCTGAACCTCCTTCTACGATCGA 3462  
QY 2726 ATCAAGTTCCTCGTGGCAGGACAGCAGCAGCACAATAATCCCCGGGAGAACCCCTTTGAT 2785  
Db 3463 ATCAAGTTCCTCGTGGCAGGACAGCAGCAGCAGCATCCCCGGGAGAACCCCTTTGAC 3522  
QY 2786 GGAGGSCATGCTTGTGTATTTCGTGGCCAAAGTGATGACATCAGATGGAACCCCTTGGTT 2845  
Db 3523 GGAGGSCATGCTTGTGTATCTCGTGGACAAGTGATGACGTGAGTGGACCCCATTTGGTT 3582  
QY 2846 GGTGTGAACATCAGTTTGTCAATAACCTCTCTTTGGATATACAATCAGCAGGCAAGAT 2905  
Db 3583 GGCCTGAACATCAGTTTGTCAATAACCTCTCTTTGGATACACAATCAGCAGGCAAGAT 3642  
QY 2906 GGACGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCCTGCGGTTTCAGCGGGCA 2965  
Db 3643 GGACGCTTTGACCTGGTGACAAATGGCGGCATCTCCATCATCTTGAGATTGCAAGGGCA 3702  
QY 2966 CTTTTCATCACACAGGACACACCTGTGGCTGCCATGGGATCGCTTCTTTGTCTATGGAA 3025  
Db 3703 CCTTTCATCACACAGGACATACCTCTGGCTACCTGGGATCGCTTCTTTGTCTATGGAA 3762  
QY 3026 ACCATCATCATGAGACATGAGGAGAAATGAGATCCCAGCTGTGACCTGAGCAATTTTGGC 3085  
Db 3763 ACCATCGTCTATGAGACACGAGGAGAAATGAGATCCCAGCAGAGACCTGAGCAACTTTGCC 3822  
QY 3086 CGCCCCAACCCAGTCTGTCTCCATCCCCACTGAGCTCTTTCGCGAGTCTGTGCGAGAG 3145  
Db 3823 CGTCCCAACCTGTGGTCTCTCCATCCCCACTGAGATCCTTTCGCGAGCTCCTGTGCTGAG 3882  
QY 3146 AAAGGCCCATTTGTCCGGAATTCAGGCTTTTGAGGAGGAAATCTCTATCTCTGGCTGC 3205  
Db 3883 AAAGGCCCATTTGTCCGAGAAATCCAGGCCCTGCAAGAGGAAATCGTCTATCGCTGGCTGC 3942  
QY 3206 AAGATGAGGCTGAGCTACCTGAGCAGCCCGACCCCTGGCTACAAATCTGTCTGAGGATC 3265  
Db 3943 AAGATGAGGTTGAGCTACCTGAGCAGCCCGCACTCCTGGCTATAAGTCTGTGCTGAGGATC 4002  
QY 3266 AGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGGTAGCGGTG 3325  
Db 4003 AGTCTCACACACCCCAACCATCCCTTCAACCTCATGAAGGTCCACCTCATGGTGGCAGTC 4062  
QY 3326 GAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCGAGCCCCAGACCTGTCTCTATTATTTCATT 3385  
Db 4063 GAGGGTCGGCTCTTTCAGGAAGTGGTTCGCGGAGCTCCTGACCTGTCTCTATTACTTTCATC 4122  
QY 3386 TGGGACAAGACAGACGTCCTACAACCAAGAGTGGTTCGGCTTTCAGAAAGCTTTTGTTC 3445  
Db 4123 TGGGACAAGACAGATGCTACAACCAAGAGTGGTTCGGACTCTCAGAAAGCTTTTGTTC 4182  
QY 3446 GTGGGTTATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTG 3505  
Db 4183 GTGGGTTATGATGATGAATCCTGCCAGATCTGATCCTGTGGGAGAAAGGACAGCGGTG 4242  
QY 3506 CTGCAGGGCTATGAAATTGACGGCTCCAGCTTGGAGGATGGAGCCTTAGACAAACATCAT 3565

Db 4243 CTGAGGGCTATGAAATTGATGCTCTCAAAGCTGGGGGCTGGAGTCTGGATAAGCACCAT 4302

Qy 3566 GCCCTCAACATTCAAAGTGGTATCCTGCACAAAGGGAATGGGGAGAACCAAGTTTGTGTCT 3625

Db 4303 GCCCTGAACATCCAGAGTGGCATCCTGCACAAAGGGAATGGAGAGAACCAAGTTTGTGTCC 4362

Qy 3626 CAGCAGCCTCTGTCTATTGGAGCATCATGGGCAATGGGCGCCGGAGAAGCATCTCCTGC 3685

Db 4363 CAGCAGCACCGGTTCATCGGAGCATCATGGGCAATGGGCAATGGGCGCAGGAGAAGTATCTCCTGC 4422

Qy 3686 CCCAGCTGCAACGGCCTTGCTGACGGCAACAAGCTCCTGSGCCCCAGTGGCCCCCACCTGT 3745

Db 4423 CCCAGCTGCAATGGTCTTGCTGATGGCAACAAGCTCCTGSGCCCCGCTGGCCCCCACCTGC 4482

Qy 3746 GGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTCCCTCT 3805

Db 4483 GGCTCTGATGGGAGTCTCTACGTGGGAGACTTCAACTACATCCGAGAATCTTCCCTCT 4542

Qy 3806 GGAATGTCACCAACATCCTAGAGCTGAGGAATAAAGATTTCAACATAGTCACAGTCCA 3865

Db 4543 GGAATGTCACCAACATCCTGGAGATG-----AGTCACAGCCCA 4581

Qy 3866 GCACACAAATACTACCTGGCCACAGACCCCATGAGTGGGCGCTCTTCCCTTCTGACAGC 3925

Db 4582 GCACACAAATACTACCTGGCTACAGACCCCATGAGTGGGCGCTCTTCCCTGCTGACACC 4641

Qy 3926 AACAGCCGCGGGTCTTTAAATCAAGTCCACTGTGGTGTGAAGGACCTTGTCAAGAAC 3985

Db 4642 AACAGCCGCGGGTCTTCAAGGTCAAGTCCACCACAGTGGTGAAGGACCTGGTGAAGAAC 4701

Qy 3986 TCTGAGTGGTTGCGGGGACAGGTGACCAGTGCCTCCCTTTTGATGACACTCGCTGCGGG 4045

Db 4702 TCCGAGTGGTAGCAGGGACTGGTGACCAGTGCCTCCCTTTTGATGATACCCGCTGCGGA 4761

Qy 4046 GATGGTGGGAAGGCCACAGAAGCCACACTCAACCAATCCCAGGGTATTACAGTGAACAAG 4105

Db 4762 GATGGTGGGAAGGCCACAGAAGCCACGCTCACTAACCCAGGGGAATTACAGTGAACAAG 4821

Qy 4106 TTTGGGTGATCTACTTTCGTGGATGGCACCATTGATCAGACGGATCGATCAGAATGGGATC 4165

Db 4822 TTTGGGTCAATTATTTCGTGGACGGCACCATGATCAGACGCTGTGATCAAAATGGGAATC 4881

Qy 4166 ATCTCCACCTGCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGTGATCT 4225

Db 4882 ATCTCCACTTTGCTGGGCTCCAATGACCTCACTCGGCCAGGCCCTCAGCTGTGACTCC 4941

Qy 4226 GTCATGGATATTTCCAGGTAAGACTGGAGTGGCCACAGACTTAGCCATCAACCCCAATG 4285

Db 4942 GTCATGGAGATTTCTCAGGTTGCGCTGGAGTGGCCACAGACTTAGCCATCAACCCCAATG 5001

Qy 4286 GACAACTCACTTTATGTCTCGACAAACAATGTGGTCTGCAAAATCTCTGAAAACCCACAG 4345

Db 5002 GATAATTCTCTATGTCTCGACAAACAATGTGGTCTGCAAAATCTCGAAAACCCACAG 5061

Qy 4346 GTGCGCATTTGTCGCGGGAGGCCCATGCACTGCCAGGTCCCTGSCATTGACCACTTCCTG 4405

Db 5062 GTACGCATTTGCTGGGAGGCCCATGCACTGCCAGGTCCCTGSCATCGACCAATTCTCTG 5121

Qy 4406 CTAAGCAAGGTGGCCATCCACGCCAACCCCTGGAGTCAGCCACCGCTTTGGCTGTTTCACAC 4465

Db 5122 CTGAGCAAGGTAGCCATCCATCGACCCCTGGAGTCAGCCCACTGCTCTGGCCGTTTTCACAC 5181

Qy 4466 AATGGGTCTCTGATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTC 4525

Db 5182 AATGGAGTCTTGTAATCGCTGAGACCGACGAGAAAAAGATCAATCGCATCAGGCAGGTC 5241

Qy 4526 ACCACTAGTGAGAGATCTCACTCGTTGCTGGGGCCCCCAGTGGTGTGACTGTAAAAAT 4585

Db 5242 ACTACAAAGTGTGAGATCTCACTGGTTGCTGGTGGCCCCCAGTGGTGTGACTGTAAAAAT 5301

Qy 4586 GATGCCAACTGTGATGTTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAGTTAAAT 4645

Db 5302 GATGCCAACTGTGACTGCTTCTCTGGAGATGATGGTTACGCCCAAGGATGCAAAGCTGAAT 5361

Qy 4646 ACCCATCTCTTGGCTGTGTGCTGATGGGAGCTCTACGTGGCCGACCTTGGGAAC 4705

Db 5362 ACCCATCTCTTGGCTGTGTGCTGACGGGAGCTCTATGTGGCCGACCTGGGAAC 5421

Qy 4706 ATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTCCCTCAACACCCAGAAACATGTATGAG 4765

Db 5422 ATCCGAATTCGATTTATCCGGAAGAACAAAGCCTTTCCCTGAACACTCAGAAACATGTACGAG 5481

Qy 4766 CTGTCTTCAACAAATTGACCAGGAGCTCTATCTGTGTTGATACCAACCGGCAAGCACCTGTAC 4825

Db 5482 CTATCTCTCCCCCATCGACCAGGAGCTGTACCTCTTTGATACCAAGTGGCAAGCATCTGTAC 5541

Qy 4826 ACCCAAAGCCTGCCACAGGAGACTACCTGTACAACCTTCACTTACACTGGGACGGCGAC 4885

Db 5542 ACTCAGAGCTACCCACAGGGGACTACCTGTACAACCTTCACTTACACAGGGACGGGAC 5601

Qy 4886 ATCACACTCATCAGACAAACAATGSCAACATGGTAAATGTCCGCGGAGACTCTACTGGG 4945

Db 5602 ATCACACATATCACCGACAAACAATGGCAACATGGTGAACGTCCGCGGAGACTCTACCCGG 5661

Qy 4946 ATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAAC 5005

Db 5662 ATGCCCTCTGGCTGGTAGTCCCAGATGGCCAGGTATACTGGGTAAACCATGGGCACCAAC 5721

Qy 5006 AGTGCACTCAAGAGTGTGACCAACAAGGACACGAGTTGGCCATGATGACATACCATGGC 5065

Db 5722 AGCGCACTCAGAAAGTGTGACCACACAAGGACACGAGCTAGCCATGATGACCTACCATGGC 5781

Qy 5066 AATTCCGGCCTTCTGGCAACCAAAAGCAATGAAACCGGATGGACAACTTTTATGAGTAC 5125

Db 5782 AACTCTGGCCTCTTGGCAACCAAAAGCAATGAAACCGGTGGACAACTTTTATGAGTAT 5841

Qy 5126 GACAGCTTTGGCGCCTGACAAATGTGACCTTCCCTACTGGCCAGTGAGCAGTTTCCGA 5185

Db 5842 GACAGTTTGTGGTGGCTGACAAACGTTGACCTTCCAACTGGCCAGGTGAGCAGTTTCCGA 5901

Qy 5186 AGTGATACAGACAGTTCAGTGCATGTCCAGGTAGAGACCTCCAGCAAGGATGATGTCAAC 5245

Db 5902 AGCGATACAGACAGCTCAGTGCACGTGCAGGTAGAGACCTCAAGCAAAAGATGACGTCAAC 5961

Qy 5246 ATAACCAACCACTGTCTGCCTCAGGCGCTTCTACACACTGTCTGCAAGACCAAGTCCGG 5305

Db 5962 ATAACCAACCACTGTCTGTTGGTGGTGGCTTCTACACCTGTCTTACAAGACCAAGTCCGG 6021

Qy 5306 AACAGCTACTACATCGGGCCGATGGCTCCTTGGCTGGCTGTCTGGCCAAACGGCATGGAG 5365

Db 6022 AACAGCTACTACATCGGGGCTGATGGCTCCTGAGGCTGCTACTAGCCAATGGCATGGAA 6081

Qy 5366 GTGGCGCTGCAGACTGAGCCCCCACTTGTGGTGGCACCCGTCAACCCCAACCGTGGGCAAG 5425

Db 6082 GTGGCTYTGCAGACTGAGCCACACCTGTCTGGTGGCACTGTCAACCCCACTGTAGGCAAG 6141

Qy 5426 AGGAATGTCAAGCTGCCCATCGACAAACGGCCCTCAACCTGGTGGAGTGGGCCAGCGCAA 5485

Db 6142 AGGAATGTCAACTGCCCATTTGATAATGGCTCAACCTGGTGGAGTGGCGGACGCGCAAG 6201

Qy 5486 GAGCAGGCTCGGGCCAGGTCACTGTCTTTGGCGCCGGCTCGGGGTGCACAAACCGAAAT 5545

Db 6202 GAGCAGGCTCGTGGCCAGGTCACTGTCTTTGGACCCCGTCTTTGGACCCCGTCTGGGGTTCACAAACCGAAAC 6261

Qy 5546 CTCCTATCTCTGGACTTTGATCGCGTAACACGACAGAGAAGATCTATGATGACCAACCGC 5605

Db 6262 CTCTTGTCTTTGGACTTTGACCGTGTCAACGCAAGAGAAGATCTACGATGACCATCGC 6321

Qy 5606 AAGTTCACCCCTTCGGATTTCTGTACGACCAAGGGGGCGGCCAGCCCTCTGGTCAACCCAGC 5665

Db 6322 AAGTTCACCCCTTCGGATCCTATATGACCAGGAGGGAGGCCAGCTTCTGCTCACTTAGC 6381

Qy 5666 ACAGGCTGAATGGTGTCAACGTTGACATATCCCTCTGGGGTTTACATTTGGTCACTCCAG 5725

Db 6382 ACAGGCTGAATGGTGTAAATGTGACCTACTCCCTGAGGTGAGGTGACATTTGCTGGAATCCAA 6441





QY 275 CGGAGCTGGGCTGGAAGAAGTAACGCCCTTACCGGACCTGTACCGGACACAT 334  
 Db 563 AGAGACATGGCTTCGGAGATCCCGTCCGCACATATGGAGCGTACCGGACAGAAATG 622  
 QY 335 GGCCTGCCCAATGCGGTACTCCATGGGGCTGGCTGTATGCCGACATGGAGGTGAC 394  
 Db 623 GGCCTTCCCGACCGGACTACTCAGTGAGTGTGGCATCAGACGCTGACACAGAAACAGAC 682  
 QY 395 ACGGTGCTGCCCTGAGCACCCCGTCCGTCTGTGGGGCGGAG--CACACGGTCAAGG 451  
 Db 683 GGCATAATGTCCCGGAGCACCGCTCCGCTCTGGGGCGGACCAACACCAATCCGGC 742  
 QY 452 CGCAGCTCCTGCTGTCCAGCCGGGCCAATTCCTCAATCTCACACTCACCGACACGAGCAT 511  
 Db 743 CGCAGTCTCCTGCTGTTCAGCAGAGCCAACTCCAACCTTACGCTCACCGACACCGAGCAT 802  
 QY 512 GAAACACTGAGACTGATCATCC-----534  
 Db 803 GAAACACTGAGATGGTCCCTCTGCAATTGCTCATCTGCCTCCTCTTCCCTGTGAC 862  
 QY 535 -----534  
 Db 863 TCCCTTACCCTCCCCCTTCCACGCAGCCAATCAGAGCCAGGAAGTTGCTAGGTAAC 922  
 QY 535 -----534  
 Db 923 AGCGGGCTCAGCCCGCGGGACTCTGAGTCGGAAGATGAGTTGGCCCAATTCATTC 982  
 QY 535 -----GGGC 538  
 Db 983 TTAGTTAAACCGGCTCAGGAAACGTCTGCACTCCGCGCGCGCTACTGCTAATGAGGGC 1042  
 QY 539 GGCCTGCAGAACCAACGCGGCTCCGGACGCCCGCGCGCTCTCGACGCCACACCC 598  
 Db 1043 TCGTTCAGAAATCACTACGGTTGCGAACGCCACCAATACCCCTGTTTCACTCGCACTCC 1102  
 QY 599 CCACACCGACACCGGGCTCCATTAACCTCCCTGAACCGGGCAACTTCACGCCGAGG 658  
 Db 1103 CCAGTCAGCACCAACACCGGCTCCATCGGTTCCCTGAGCCGCGAGCAACTACACCCAGCGC 1162  
 QY 659 AGCAACCCAGCCCGCCCGCCACCGACCACTCGCTCTCCGGAGAGCCCTGCGCGGCGC 718  
 Db 1163 AGTAACCCGAGCCCGCACCCACCGATAGTTCAAGTCCCAATGAGGGTCCGAC--AAGT 1219  
 QY 719 GCCCAGGAGCCTGCCCAACCGCCAGGAGAACTGGTGTCTCAACAGCAACACTCCCTGGAG 778  
 Db 1220 GCCCAGGACTCCAGCAGCGCCAGGACAACTGGCTTCTCAACAGCAACCTTCTCTAGAG 1279  
 QY 779 ACCAGGAACCTAGGCAAGCAGCCATTCCTAGGACATTCAGGACAACTCATTTAGATG 838  
 Db 1280 ACCAGAAACATAGCAAAAGCAGACATTCCTAGAGACTTTGAGGACAACTTTATAGATG 1339  
 QY 839 GACATTCCTCGCGCCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACTGACGTCCAGC 958  
 Db 1340 GACATACTGGCGACAGCTCGCGCGACGCGCTTACACTGATGGACACTTCTTTTCAAG 1399  
 QY 899 CCTGGAGGCACCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACTGACGTCCAGC 958  
 Db 1400 CCTGGCGGACCTCGCCTTTGTACTGCACCACTTACCTGGTTACCTTTGACGTCCAGC 1459  
 QY 959 ACAGTGTACTCTCCTCCGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGCGCGCTTT 1018  
 Db 1460 ACGGTGTACTCTCCTCCGCCCCGCCCCCTGCAAGAAACACCTTCTACGGCCTGCTTT 1519  
 QY 1019 AACCTCAAGAAAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCCATCTC 1078  
 Db 1520 AGCCTAAAGAAACCCCTATAAGCACTGCACTGGAATGTGCTGCCCTCAGTGCCATCCTC 1579  
 QY 1079 ATCTCAGCCACTCTGCTCATCTGCTGGCACTATTGTTGGCCATGACCTGTTTGGCCCTA 1138  
 Db 1580 ATCTCCGTACCCCTGGTGTCTGTTGGCACTATTATTGCCATGACCTCTTTGGACTC 1639  
 QY 1139 AACTGGCACCTGACGCCGATGGAGGGGACATGTATGATATCAGGAGGACACAGCCAGC 1198

Db 1640 AACTGGCATTTGCAGCCGGTGCAGAGGCAGATATATCACTGACGGAGACAAACACAGT 1699  
 QY 1199 AGTTGGCTGTGCCAACCGACGTCTCCTATATACCCCTCAGGGGACACTGGCTTAGAGACC 1258  
 Db 1700 GGCCTACATCTGCCACAGATCTGGGGCTGCCACCTCTGGGTAAACACAGGTCTAGAGTTT 1759  
 QY 1259 CCTGACAGGAAGGCAAGGAACCAAGAAAGGAAAGCCAGTAGTTTCTTTCCAGAGGAC 1318  
 Db 1760 CCAGATCGAGGAAGCAGAGAT-----GATGGTAAACTAGATGGCTTCTTCCCGAGGAC 1813  
 QY 1319 AGTTTCATAGATTCTGGAGAAATGTATGTGGAAAGGCGAGCCTCCAGAAAGATTCTCCT 1378  
 Db 1814 AGCTTCATAGACATGGCGGAGATAGACGTGGCGGCAAAAGTGGCTCAGCTGATCCCGCT 1873  
 QY 1379 GGCACCTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAAATTCATGTG 1438  
 Db 1874 GGAATCTTTGGAGGTGCGAGGTCTTTATTGATCACCCCATGTACCTCAAAATTCATGTG 1933  
 QY 1439 TCTCTGGAAAGGCGAGCCCTGGTTGGCAATTTATGCGAGAAAAGGCCCTCCCTCTTCCAT 1498  
 Db 1934 TCCCTTAGTAAAGATGCTTGGTGGGATCTACGGGAGGAGGACTGCCACCGTCCAC 1993  
 QY 1499 ACACAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCCTAAACCCAGGAGCGCG 1558  
 Db 1994 ACACAGTTTGACTTTGTGAGCTGTTGGATGGTCTGCTGGCTTCTGTGCAAGGCTTGCCA 2053  
 QY 1559 AGCCTAGAGGGGAC-----CCCGCGCCAGTCTCGGGGAACTGTGCCCCCTCCAGCCAT 1612  
 Db 2054 GGTCTTGATGGTCCGCCCTTTCCAGCCCGCAGCGGAGCCTGGTGCCTATAACCACTCAT 2113  
 QY 1613 GAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTACAATGAC 1672  
 Db 2114 GATACAGGATGTATACAGTACATGGACTCAGGAATCTGGCACTTGGCTGTATACAATGAT 2173  
 QY 1673 GGAAGGAGTCAGAAAGTGGTTTCTTTCTACCACTGCCATTGAGTCGGTGGATAACTGC 1732  
 Db 2174 GGAAGGAGACTGAGCAGGTGTCTTTCTAACAACTGCAATCGACTTATGATGACTGT 2233  
 QY 1733 CCCAGCAACTGCTATGGCAATGGTACTGCTCTCTCTGGGACCTGCCACTGCTTCTGGGT 1792  
 Db 2234 CCTAGTAACTGTTTGGAAATGGAGACTGTGTCTGGAACCTGCCATTTGTTTCCAGGC 2293  
 QY 1793 TTCCTGGCCCCGACTGTGGCAGAGCCTCTTCCCTGCTCTGTAGCGGAAATGGCCAA 1852  
 Db 2294 TTCAGAGGACCCGACTGCAGCAGGCGCATCATGTCTGTCTGTGTCAGTGGAAACGGTCAG 2353  
 QY 1853 TACATGAAAGGCAGATGCTTGTGCCACAGTGGCTGGAAGCGCTGAGTCGATGTGCC 1912  
 Db 2354 TACCTGAAAGGACGCTGCATGTGTACAGCGGCTGGAAGGGTCTGAATGTGATGTTCT 2413  
 QY 1913 ACCAACCCAGTGTATCGATGTGGCTGCGCCTGCAGCAACCATGGCACCTGCATCAGCGGCACCTGC 1972  
 Db 2414 ACAACCCAGTGCATCGACATCACCTGCAGTGGACATGGAACATGCATGTGGGAACCTGC 2473  
 QY 1973 ATCTGCAACCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCA 2032  
 Db 2474 ATCTGCAATCCTAGTACAAGGAGAAACTGCGAAGAGTGGACTGTCTGGACCCCAACC 2533  
 QY 2033 TGTTCAAGCCCGGGTGTCTGCTGAGAGGCGAATGCCATTGCTTTTGGGATGGGAGGC 2092  
 Db 2534 TGTTCAAGGAGAGTGTGCTGCTGCGGTTCGGGAGAGTGCCACTGTTTGTGGGATGGGAGGT 2593  
 QY 2093 ACCAACTGCGAGACCCCGGCGCACATGCTTAGACCAGTGTTCAGGCCACGGAACCTTC 2152  
 Db 2594 CCTGGATGTGAGATCCTCGTGGCTCTGTATGGAGCAGTGTCTGGACACGGGAGTTTC 2653  
 QY 2153 CTCCCGGACACCGGGCTTTGCAGCTGTACCCCAAGCTGGACTGGACACGACTGTTCTATC 2212  
 Db 2654 TTGGCTGATACAAACACCTGCAACTGTGACCACAACTGGACCGGTCTGCTCCACA 2713  
 QY 2213 GAGATCTGTGCTCGGACTGTGGTGGCCATGGCGTGTGCGTAGGGGACCTGCCGCTGC 2272

Db 2714 GAGCTGTGCGCAGCGGACTGTGGTGGACATGGCATCTGTGTGGCGGGCAGCTGCCCGGTGC 2773  
QY 2273 GAGGATGCTGGATGGGGGCGAGCCCTGCGACCGCGGCCCTGCCACCCGCGCTGTGCCGAG 2332  
Db 2774 GACGAGGCTGGATGGGCAACCGGTGCGAGCAACGGGCTTGTCACCTCGCTGCAGCGAA 2833  
QY 2333 CATGGGACCTGCCGCGACGCGCAAGTCCGAGTGCAGCCCTGCTGGAATGGCGAACACTGC 2392  
Db 2834 CACGGGACCTGCAAGGACGGAAGTCCGAATGCAGCCCGGGATGGAATGGAGAACACTGC 2893  
QY 2393 ACCATCGCTCACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCCCTGGGTTGTGCAATGGC 2452  
Db 2894 ACTATC-----GAGGGCTGCCCTGGTCTGTGCAATGGC 2926  
QY 2453 AACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGA 2512  
Db 2927 AATGGCAGATGCACGCTGGGTAATAATGGCTGGTACTGCGTTGTCAACTGGGCTGGCGG 2986  
QY 2513 GGAGCTGGTGTGACACTTCCATGGAGACTGCCCTGCGGTGACAGCAAAAGACAATGATGGA 2572  
Db 2987 GGAGCAGGCTGCGACACATCAATGGAGACGGCCTGCAGTGCAGCGCAAGGACAATGATGGA 3046  
QY 2573 GATGGCCTGGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAAC 2632  
Db 3047 GATGGGTTGACGGACTGCATGGACCCGGATTGCTGCTTGCAAGGCATCATGTACACAACA 3106  
QY 2633 CCGCTGTGCTTGGCTCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCTCTGTG 2692  
Db 3107 TCTCTGTGTGGGCTCTCCCGATCCCTCTGGATATCATTCAGAAACCCAAATCTCTTCC 3166  
QY 2693 TCACAGCAGAACCTACACTCCTCTATGACCGCATCAAGTTCCTCGTGGCAGGGACAGC 2752  
Db 3167 AGTCTGAGCACCCTGCAGTCCCTTCTATCAGCGCATCCATTTCTTGGTGGCCGGGACAGC 3226  
QY 2753 ACGCACATAATCCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGC 2812  
Db 3227 ACCCATGTATACCTGATGTCAATCCCTTTGATGGCAATTCATGCTGCGTTATCCGGGGA 3286  
QY 2813 CAAGTGATGACATCAGATGGAACCCCCCTGGTGGTGTGAACATCAGTTTTGTCAATAAC 2872  
Db 3287 CAAGTGGTGACTTCTGATGGAACCTCTTGGTGGCGTAATATCAGCTTCAATTACAAA 3346  
QY 2873 CCTCTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGC 2932  
Db 3347 CCGGCCTATGGGTATACAATCACAGACAGGACGGCAGCTTTGACCTGGTAAAGCAACGGT 3406  
QY 2933 GGCATCTCCATCATCTCGGTTGAGCGGGCACCTTTATACACACAGGAGCACACCCCTG 2992  
Db 3407 GCGTGGCCATGCGCCTGCGGTTTGAGCGTGTCTCCCTTTATCACTCAGGAGCACACCTTG 3466  
QY 2993 TGGCTGCCATGGGATCGCTTCTTTGTATGGAACCATCATCATGAGACATGAGGAGAAAT 3052  
Db 3467 TGGCTGCCGTGGGGCGGTTCTTCGTATGATACAATTTGATGCGACACGAGGTC AAC 3526  
QY 3053 GAGATTCACAGCTGTACCTGAGCAATTTTCCCGCCQCCAAACCCAGTCTCTCTCCATCC 3112  
Db 3527 GACATCCCCAGCTGCGACCTGAGTAGCTTCACCCGTCCCATGCCCCATCGTCTCGCTGCT 3586  
QY 3113 CCACTGACGTCTTCGCCAGCTCCTGTGACAGAGAAAGGCCCATTTGTCCCGGAAATTCAG 3172  
Db 3587 CCTTTGACAGCGGTTTGACAGGACCTGCCCTGAGAGGGGCATCGTCGTTCCAGAGATTGAG 3646  
QY 3173 GCTTTGACAGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGC 3232  
Db 3647 ACCCTCCAAGAGAAGTGGGATCCCAAGTACAGATATGCGCTTGGGCTACTTGAGCAGC 3706  
QY 3233 CGGACCCCTGGGTACAAATCTGTCTGAGGATCAGCCCTCACCCACCCGACCATCCCCCTTC 3292  
Db 3707 CGGACTTCTGGGTATAAGTCTCTTCTGGAATCACCCCTCACCCACTCCACCATCCCCCTTC 3766  
QY 3293 AACCTCATGAGGTGCACCTCATGTGACGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTC 3352  
Db 3767 AGCCTGATGAAGGTGCACCTTATGGTGGCCGTTGAAGGAAGACTTTTCCGGAAGTGGTTT 3826

QY 3353 GCTGCAGCCCCCAGACCTGTCTCTATTATTTTCAATTTGGGACAAAGACAGACGCTCTACAACCAG 3412  
Db 3827 TCAGCTGCTCCAAACCTCTCTTATGACTTGTTTGGGACAAAACCGACGCTCTACAGCCAG 3886  
QY 3413 AAGGTGTTTGGGCTTTCAGAAGCCCTTTGTTTCCGTGGGTTATGAATATGAATCCTGCCCA 3472  
Db 3887 AAAGTATATGGGCTGTCTGAGGCTTTTGTGTCCGTGGGTTTGTAGTATGAATCCTGTCCA 3946  
QY 3473 GATCTAATCCTGTGGGAAAAAAGAACACAGTCTGCAGGGCTATGAAATGACGCGTCC 3532  
Db 3947 GATCTAATCCTGTGGGAGAAAGCGGACGSGCTGTTCTGCAGGGATACGAGACCACCTGCGTCA 4006  
QY 3533 AAGCTTGGAGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTG 3592  
Db 4007 AATCTTGGTGGCTGAGTGTGGACAAACACCATGCGCTGAACATTTCAAAGTGGTATCCTT 4066  
QY 3593 CACAAAGGGGAATGGGGAGAACCCAGTTTGTGTCTCAGCAGCCCTCCTGTCTATTTGGGAGCATC 3652  
Db 4067 CACAAGGGCAACGGTGAGAACATCTTTCATTTCCAGCAGCCTCCAGTCCATAGGAAGCATC 4126  
QY 3653 ATGGGCAATGGGCGCCGGAGAAAGCATCTCCTGCCCCCAGCTGCACGGCCCTTGTGACGGC 3712  
Db 4127 ATGGGTAATGGACGGAGGCGCAGTATCTCTGTCCGAGCTGCACGGGCTGGCGGATGGA 4186  
QY 3713 AACAAAGCTCCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGCACGGGAGCCTCTATGTGGGT 3772  
Db 4187 AACAAAGCTGCTCGACCTGTAGCTTTAGCTTGGGTCGGATGGAGTCTGTATGTGGGA 4246  
QY 3773 GATTTCAACTACATTAGAAGGATCTTCCCTCTGGAATGTCAACCAACATCCTAGAGCTG 3832  
Db 4247 GATTTCAACTACGTCGTCGCATCTTCAACCGGGAATGTCAACCGGCTCCTAGAGCTC 4306  
QY 3833 AGGAATAAAGATTTACAGACATAGTACAGTCCAGCACACAAATACTACCTGGCCACAGAC 3892  
Db 4307 AGCA-----ACAGCCCTGCACACAAATATTACCTGGCCACCAGT 4345  
QY 3893 CCCATGAGTGGGSCCGTCTTCTCTTCTGACAGCAACAGCCGGGCTTTTAAATCAAG 3952  
Db 4346 CCTGTTTCAGGTTGGCTGTACCTCTCGGACACACAGCAGCAGGAAAGGTGTTCAAAGTCAAG 4405  
QY 3953 TCCACTGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTTGGGGGACAGGTGAC 4012  
Db 4406 TCTCTGTATGCAATAAAGATGTGGCCAAAACCTGGAGCTGTTGGCGGAACAGGAGAT 4465  
QY 4013 CAGTGCCTCCCTTTGATGACACTCGCTGCGGGGATGTTGGGAAGGCCACAGAAAGCCACA 4072  
Db 4466 CAGTGTCTGCCATATGACGAGACTCGATGTGGGATGGAGGCAAAAGTGTGGAGGCCACA 4525  
QY 4073 CTCACCAATCCCGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGTTGATGGC 4132  
Db 4526 CTCACCAACCCAGAGGCATCACCGTGGATAAGTATGTTGTGATCTCTTTTGTGATGGC 4585  
QY 4133 ACCATGATCAGACGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGAT 4192  
Db 4586 ACCATGATCCCGCCGATCGATCAGAACGGCATCATATCCACCCCTGCTGGGCTTCAACGAC 4645  
QY 4193 CTCACATCAGCCCGCCACTCAGCTGTGATTCTGTCTATGGATATTTTCCAGGTAAGACTG 4252  
Db 4646 TTGACTTGGGCTCGACCCCTTAGCTGTGACTCGGTGATGGACATCTCGCAGGTGCGCCTG 4705  
QY 4253 GAGTGGCCACAGACTTAGCCATCAACCCAAATGGACAACTCATTATGTCTCGACAAC 4312  
Db 4706 GAGTGGCCGACTGACCTCGCTGTGAGCCCCATGGACAACTCTCTTTATGTGTGGACAAC 4765  
QY 4313 AATGTGGTCTTGCAAAATCTCTGAAAAACCCAGGTGCGCATTTGTGCGGGAGGCCCATG 4372  
Db 4766 AATGTAGTCTTGCAGATCTCAGAAAAACCATCAGGTGAGGATTGTAGCAGGACGCCCCATG 4825  
QY 4373 CACTGCCAGGTCCCTGGCATTGACCACTTCTCTGTAAGCAAGGTGGCCATCCACGCAACC 4432  
Db 4826 CACTGCCAGGTCCCGGCTCTGGACCACTTCTTAGTCAGCAAAATCGCCATCCATGCCACC 4885



Db 7046 GCAAATACCAACCCGCTATGCCTATGAATATGATGCAGATGGTCAGCTCCAGGTAGTGTC 7105

QY 6653 ATCAATGACAAGCCACTCTGGCTACAGCTACGACCTCAATGGGAACCTGCACCTACTG 6712

Db 7106 ATCAATGATAAACCTCTATGGCGTACAGTTATGACCTTAATGGCAATTTGCACCTCCTT 7165

QY 6713 AGCCCTGGGAACAGTGCACGGCTCACACCACACTACGGTATGACATCCGCGACCGCATCACT 6772

Db 7166 AGCCCTGGTAATAGCGCCAGACTTACACCACACTACGCTACGACATTCGGGACCGCATTACT 7225

QY 6773 CGGCTGGGTGACGTGCAATACAGATGGATGAGGATGGCTTCCTGAGGCAGCGGGCGGT 6832

Db 7226 CGTCTAGGTGATGTCAGTACCGTTTAGATGAGGATGGATTCTCTCAGGCAAGAGGCAAT 7285

QY 6833 GATATCTTTGAGTACAACACTCAGCTGGCCTGCTCATCAAGGCTTACAACCGGGCTGGCAGC 6892

Db 7286 GATTTCTTTGAGTATAACTCTGCTGGGTGCTTGTTTAAACCTACAACAAAGTTAACGGA 7345

QY 6893 TGGAGTGTACAGGTACCGCTACGATGGCCTGGGGCGGGCGGTGTCCAGCAAGAGACGCCAC 6952

Db 7346 TGGACTATCAAGTACCGCTATGATGGCTTGGGCAGAAAGATGTCCAGTCGAAAGTACCCAA 7405

QY 6953 AGCCACACCTGCAGTTCTTCTATGAGACCTGACCAACCCCAAGGTCAACCCACCTG 7012

Db 7406 GGCCACCATCTGCAATTCTTCTATGCAGACTTGTCCAGTCCCAACAGAGTCAACACATG 7465

QY 7013 TACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGAAGGACACCTCTTT 7072

Db 7466 TACAATCACTCAAGCTCAGAAATCACTCACTACTATGACCTTCAAGGCCACCTTTT 7525

QY 7073 GCCATGGAGTGAGCAGTGGTGATGAGTTTACATAGCTTTGTGACAAACATCGGGACCCCT 7132

Db 7526 GCCATGGAGTTGAGCAGTGGGGATGAGTTCTATGTGGCATGGGACAAACATTTGTTACTCA 7585

QY 7133 CTTGCTGTCTTTAGTGGAAACAGTTTGTATGATCAAGCAAACTCTGTACACAGCCTATGGG 7192

Db 7586 CTGGCTGTATTAGTGGCGTGGTCTCATGATCAACAGATTTCTCCACACAGCTTTTGGG 7645

QY 7193 GAGATCTACATGGATACCAACCCCAACTTTTCAATCATCATAGGCTACCATGTTGGCCTC 7252

Db 7646 GAAGTATATTAGACTCCAACCCCAAGCTTCCAGCTAGTAATAGGCTACCAAGGCGGCTT 7705

QY 7253 TATGATCCACTCACCAAGCTTGTTCACATGGCGCGGCGAGATTATGATGTGCTGGCCGGA 7312

Db 7706 TATGAGCCCTTACAAAGCTTGTCCATATGGCGCGCAGAGATTATGATGTTCTTGACGC 7765

QY 7313 CGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCT 7372

Db 7766 CGGTGGACCACTCTGACCATGATATCCGGAAGCGGCTCAACAGCGACAAACATTTGCGCC 7825

QY 7373 TTTAATCTCTATATGTTCAAAAACAAACCCCATCAGCAACTCCAGGACATCAAGTGC 7432

Db 7826 TTCAACCTCTACATGTTTAAAGAACAAACCCCTTGAGCAACAGCCCAAGAAACCAAGTGC 7885

QY 7433 TTCAATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAAACGTGATC 7492

Db 7886 TACATGACAGATGTTAAACAGCTGGCTGGTGACGTTTCGGTTTCCAGCTATACAAACGTATC 7945

QY 7493 CCTGGTTATCCAAACCCAGACATGGATGGCATGGAACCCCTCCTACGAGCTCATCCACACA 7552

Db 7946 CCTGGCTACCGCAAGCCTGTCAACAGATGCCATGGAGCCTTCGTACGAGCTCGTTTCACACC 8005

QY 7553 CAGATGAAAACGAGGAGTGGGACACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTA 7612

Db 8006 CAGATAAAGACTCAGGAGTGGGATAGCAACCAAGTCTGTTCTAGGGGTGAGTGTGAAGTC 8065

QY 7613 CAGAAGCAGCTCAAGCCCTTTGTACCTTAGAACGGTTTGAACAGCTCTATGGCTCCACA 7672

Db 8066 CAGAGGCAGCTGAAGTCTTTTGTCCGGTGGAGCGGTTTGGTCAGATCTACAGTGCAGT 8125

QY 7673 ATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAGTTTGTGATCCAGCGGCTCAGTCTTT 7732

Db 8126 GATTTCTGGATGCTCTCCAAACCCCTCTCCACACACTTTTGTGCAACAGGGACTTCTCTCTTC 8185

QY 7733 GGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCACAGACATCATCAGTGTG 7792

Db 8186 GGGAAAGGTGTGAAAGTCCGCTATCCGTGAAGGCCGCTTGAAGGGACATCATCAGCCTG 8245

QY 7793 GCCAATGAGATGGGCGAAGGTTTGTCTGCCATCTTGAACCATGCCCCACTACCTAGAGAAC 7852

Db 8246 GCTAATGAAGATGGCCGAAGGATCGCCGAGTCTCTAGACAAGGCTAGTACCTGCAGGAC 8305

QY 7853 CTGCACCTTACCATTTGATGGGTGGATACCCATTACTTTTGTGAACACAGGACCTTTCAGAA 7912

Db 8306 TTGCATTTACCATTTGAGGCTTGGACACGCACACTACTTTTGTCAAAATCAGGTCTGGTAGAG 8365

QY 7913 GGTGACCTGGCCATCTGGGCTCAGTGGGGGGCGGCGAACCCCTGAGAGATGGGTCAAC 7972

Db 8366 GGTGATCTATCTCTGTGGTATGACAGTAGGCCAGCGTACGCTGGAGACAGGAGTCAAC 8425

QY 7973 GTCACCTGTGCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATC 8032

Db 8426 GTAACAGTCTCGCAGGTCAACATGCTCTCGGAGGACGTTTCGCGCGGCATACCCGACATC 8485

QY 8033 CAGCTCCAGTACGGGGCACTGTGCTTGAACACACAGCTACGGGACAACTTGGATGAGGAG 8092

Db 8486 CAGATGCAATACGGTACTCTAAGCCTTAATGTACGCTACGGCAGCAGTGTGGATGAAGAA 8545

QY 8093 AAGGCACGGGTCTCTGGAGCTGGCCCGCAGAGAGCCGCTGCGCCAAAGCTGGGCCCGGAG 8152

Db 8546 AAGGTACAGAGTCTGTGAGCTTGACGACAGAGGGCGGTTGCAACCGCATGGGCCACGAG 8605

QY 8153 CAGCAGAGACTGCGGGAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8212

Db 8606 CGCCACAGACTACGGCAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8665

QY 8213 CAGGTGCTGAGCAGAGGGCGGTTGCAAGGCTACGACGCGCTTTTTCGTGATCTCTGTGAG 8272

Db 8666 CAGCTCCTAAGCTGTGACGCTGTACAGGCTACGAAGGCTTCTACATAGTATCAGTGCAC 8725

QY 8273 CAGTACCCAGAACTGTACAGACAGCGGCCCAACAACTCCACTTTCATGACAGACAGGAGATG 8332

Db 8726 CAGTTCCAGAGTTGACTGACAAACATAAATAACGTCCTTCTTCTGGCAGACAGACTGAGATG 8785

QY 8333 GGCCGGAGGTGACA 8346

Db 8786 GGACGCAGGTGACA 8799

RESULT 7  
HSM806812  
LOCUS HSM806812 8993 bp mRNA linear PRI 28-AUG-2003  
DEFINITION Homo sapiens mRNA; cDNA DKFZp686K11107 (from clone DKFZp686K11107).  
ACCESSION BX640737  
VERSION BX640737.1 GI:34364828  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 8993)  
AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,  
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and  
Wiemann,S.  
CONSRMT The German Human cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center,  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp686K11107) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further

information about the clone and the sequencing project is available  
at <http://mips.gsf.de/proj/cdna/>.

FEATURES

```
source
1..8993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q14.1"
/clone="DKFZp686K11107"
/tissue_type="human fetal kidney"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="fetal"
1..8993
/gene="DKFZp686K11107"
<1..4200
/gene="DKFZp686K11107"
/note="hypothetical protein, N-terminus elongated,
differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAE45850.1"
/db_xref="GI:34364829"
/translation="RIDQNGIISTLLGNSDLTSARPLSCDSVMDISQVHLEWPTDLAI
NPMDSNLYLVNDNVLIQISENHQVRIIVAGRPMHCQVPGIDHFLLSKVAIHATLESAlA
LAVSHNGVLYIAETDEKKINRIRQVTSGEISLVAGAPSGCCCKNDANDCDFSGDDGY
AKDAKLNTPSSILAVCADGELYVADLGNIRIRFKNKPFLNTQMYELSSPIDQELYL
FDTTKHLTYQSLPTGDLNFTYTGDDITLITDNNGMNVNRDSTGMPPLWLVPD
GVYVVTMTGNSALKSVTTQGHELAMTYHGNSGLLATKSNENGWTFYEYDSFGRLT
NVTFTPTQVSSPFRSDTSSVHVQVETSSKDDVTITNLSASGAFYLLQDQVRNSYII
GADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRVNLTLPIDNGLNLVEWRQKEQA
RGQVTVFGRRLRVHNRNLLSLDFDRVTEKIDYDHRKFTLRILYDQAGRPSLWSPSS
RLNGVNVTVSPGVIAGIQIGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLL
LHSQROYIFEFDKNDRLSSVMPNVARQTELTIRSVGYRNIYQPPEGNASVIOFTE
DGHLLHTFYLGTGRVRYKYKGLSKLAETLYDTKVSFTYDETAGMLKTINLQNEGFT
CTIRYRQIGPLIDRQIPRFTTEGMVNARFDYNDNSFRVTSMOAVINETPLPIDLYRY
DDVSGKTEQFGKFGVIVYDINQIITTAVMTHTKHFDAYGRMEVQVEIFRSLMYMTV
QYDNMGRVVKELKVGPPYANTTRYSEYDADGQLQTVSINDKPLWRSYDLNGLNLHL
SPGNSARLTPLRYDIRDRIIRLGDVQVKMEDGDFLRQGGDIFEYNSAGLLIKAYNEA
GWSVRYRYDGLGRVRVSSKSHSHLOFFYADLTNPTKVTHLYNHSSEITSLYVDLQ
GHLFAMELSSGDEFYIACDNIGTPLAVSGTGLMIKQILYTAIGEIMYDTPNFQIIL
GYHGLYDPLTLKLVHMRDRDYDLVAGRWTS PDHELWKHLSSNMPFNLYMFKNNPI
SNSQDIKCFMTDVNSWLLTFGFQHLNVI PGYPKPDMDAMEPSYELIHTQMKTQWDNS
KSLGVQCEVQKLFATLRFDFDLYGSTITSCQAPTKKFPASSGSVFGKGVKRAL
KDGRTVTDIIISVANEDGRRVAAI LNHAHYLENLHFTIDGVDTHYFVKPSPSEGDLAIL
KLGGRRVTTLENGVNVTVSQINTVLNGLRTRYTDIQLQYALCLNTRYGTTLDEERARV
LELARQRAVRQAWAREQORLREGEEGLRAWTEGEKQQLVSTGRVQGYDGFVIVSBEQY
PELSDSANNTHEMRQSEMGR"
8953..8958
/gene="DKFZp686K11107"
8973
/gene="DKFZp686K11107"

polyA_signal
polyA_site

ORIGIN
```

```
Query Match 50.2%; Score 4197.2; DB 9; Length 8993;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 4202; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 4145 CGCATCGATCAGAATGGGATCATCTCCACCCTGTCGGCTCTAATGATCTCACATCAGCC 4204
Db 1 CGCATCGATCAGAATGGGATCATCTCCACCCTGTCGGCTCTAATGATCTCACATCAGCC 60

Qy 4205 CGGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCAGGTAAGACTGGAGTGGCCCCACA 4264
Db 61 CGGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCAGGTTACCTGGAGTGGCCCCACA 120

Qy 4265 GACTTAGCCATCAACCCAAATGGACAACACTCACTTATGTCTCTGACAAACAATGTGGTCCTG 4324
Db 121 GACTTAGCCATCAACCCAAATGGACAACACTCACTTATGTCTCTGACAAACAATGTGGTCCTG 180

Qy 4325 CAAATCTCTGAAAACCAACAGGTGCGCATTTGTCGGGGAGGCCCATGCACTGCCAGGTC 4384
Db 181 CAAATCTCTGAAAACCAACAGGTGCGCATTTGTCGGGGAGGCCCATGCACTGCCAGGTC 240
```

```
Qy 4385 CTGGGCATTGACCACCTTCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCC 4444
Db 241 CTGGGCATTGACCACCTTCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCC 300

Qy 4445 ACCGCTTTGGCTGTTTACACAAATGGGGTCTCTGTATATTGCTGAGACTGATGAGAAAAAG 4504
Db 301 ATCGCTTTGGCTGTTTACACAAATGGGGTCTCTGTATATTGCTGAGACTGATGAGAAAAAG 360

Qy 4505 ATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCCCCC 4564
Db 361 ATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCCCCC 420

Qy 4565 AGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTAT 4624
Db 421 AGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTAT 480

Qy 4625 GCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGTGCTGATGGGAGCTC 4684
Db 481 GCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGTGCTGATGGGAGCTC 540

Qy 4685 TAGCTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTCTC 4744
Db 541 TAGCTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTCTC 600

Qy 4745 AACACCCAGAACATGTATGAGCTGTCTTCCACCAATTGACCAGGAGCTCTATCTGTTGAT 4804
Db 601 AACACCCAGAACATGTATGAGCTGTCTTCCACCAATTGACCAGGAGCTCTATCTGTTGAT 660

Qy 4805 ACCACCGGCAAGCACCTGTACACCCAAAGCCTTCCACAGGAGACTACTCTGTACAACTTC 4864
Db 661 ACCACCGGCAAGCACCTGTACACCCAAAGCCTTCCACAGGAGACTACTCTGTACAACTTC 720

Qy 4865 ACCTACACTGGGGACGGCGACATCACACTCATACAGACAACAATGGCAACATGGTAAAT 4924
Db 721 ACCTACACTGGGGACGGCGACATCACACTCATACAGACAACAATGGCAACATGGTAAAT 780

Qy 4925 GTCCGCCGAGACTCTACTGGGATGCCCTCTGCTGGTGGTCCCAGATGGCCAGGTGTAC 4984
Db 781 GTCCGCCGAGACTCTACTGGGATGCCCTCTGCTGGTGGTCCCAGATGGCCAGGTGTAC 840

Qy 4985 TGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCAACACAGACACGAGTTG 5044
Db 841 TGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCAACACAGACACGAGTTG 900

Qy 5045 GCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAACGGA 5104
Db 901 GCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAACGGA 960

Qy 5105 TGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCTACT 5164
Db 961 TGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCTACT 1020

Qy 5165 GCGCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAAGTGCATGTCCAGGTAGAGACC 5224
Db 1021 GCGCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAAGTGCATGTCCAGGTAGAGACC 1080

Qy 5225 TCCAGCAAGGATGATGTACCATAAACCACCACTGTCTGCTCAGGCGCTTCTACACA 5284
Db 1081 TCCAGCAAGGATGATGTACCATAAACCACCACTGTCTGCTCAGGCGCTTCTACACA 1140

Qy 5285 CTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGCTCTTGGCGCTG 5344
Db 1141 CTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGCTCTTGGCGCTG 1200

Qy 5345 CTGCTGGCCAAACGGCATGGAGGTGGCGCTGAGACTGAGCCCCACCTTGGCTGGCACC 5404
Db 1201 CTGCTGGCCAAACGGCATGGAGGTGGCGCTGAGACTGAGCCCCACCTTGGCTGGCACC 1260

Qy 5405 GTCAACCCCAACCGTGGGCAAGAGGAATGTACGCTGCCCCATCGACAACGGCCTCAACCTG 5464
Db 1261 GTCAACCCCAACCGTGGGCAAGAGGAATGTACGCTGCCCCATCGACAACGGCCTCAACCTG 1320
```

QY 5465 GTGAGTGGCGCCAGCGCAAAAGAGCAGGCTCGGGGCCAGGTCACGTCTCTTTGGGCGCCGG 5524  
Dd |||||  
Dd 1321 GTGAGTGGCGCCAGCGCAAAAGAGCAGGCTCGGGGCCAGGTCACGTCTCTTTGGGCGCCGG 1380  
QY 5525 CTGCGGTTGCACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAAACACGCACAGAG 5584  
Dd |||||  
Dd 1381 CTGCGGTTTCAACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAAACACGCACAGAG 1440  
QY 5585 AAGATCTATGATGACCACCGCAAGTTTCACCCCTCGGATTTCTGTACGACGCGGGCGG 5644  
Dd |||||  
Dd 1441 AAGATCTATGATGACCACCGCAAGTTTCACCCCTCGGATTTCTGTACGACGCGGGCGG 1500  
QY 5645 CCCAGCCTCTGGTCAACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATATCCTCTGG 5704  
Dd |||||  
Dd 1501 CCCAGCCTCTGGTCAACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATATCCTCTGG 1560  
QY 5705 GGTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCAGCG 5764  
Dd |||||  
Dd 1561 GGTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCAGCG 1620  
QY 5765 GGCCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATATTAGAG 5824  
Dd |||||  
Dd 1621 GGCCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATATTAGAG 1680  
QY 5825 AAGTCCATGGTGTGCTACTACACAGCCAGCGCAGTATATCTTTGAGTTGACAAGAAT 5884  
Dd |||||  
Dd 1681 AAGTCCATGGTGTGCTACTACACAGCCAGGCGAGTATATCTTTGAGTTGACAAGAAT 1740  
QY 5885 GACCGCCTCTCTTCTGTGACGATGCCAAACGTGGCGGCGAGACACTAGAGACCATCCGC 5944  
Dd |||||  
Dd 1741 GACCGCCTCTCTTCTGTGACGATGCCAAACGTGGCGGCGAGACACTAGAGACCATCCGC 1800  
QY 5945 TCAGTGGGCTACTACAGAAACATCTATACGCCCCCTGAGGGCAATGCCTCAGTCATACAG 6004  
Dd |||||  
Dd 1801 TCAGTGGGCTACTACAGAAACATCTATACGCCCCCTGAGGGCAATGCCTCAGTCATACAG 1860  
QY 6005 GACTTCACTGAGGATGGGCACCTCCTTCAACACCTTCTACCTGGGACTGGCCGAGGGTG 6064  
Dd |||||  
Dd 1861 GACTTCACTGAGGATGGGCACCTCCTTCAACACCTTCTACCTGGGACTGGCCGAGGGTG 1920  
QY 6065 ATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTC 6124  
Dd |||||  
Dd 1921 ATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTC 1980  
QY 6125 AGTTTCACCTATGACGAGACGGCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGSC 6184  
Dd |||||  
Dd 1981 AGTTTCACCTATGACGAGACGGCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGSC 2040  
QY 6185 TTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCCCTGATTGACCGACAGATCTTCCGC 6244  
Dd |||||  
Dd 2041 TTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCCCTGATTGACCGACAGATCTTCCGC 2100  
QY 6245 TTCACCTGAGGAAGGCATGGTCAACGCCCCGTTTGACTACAACCTATGACAACAGCTTCCGG 6304  
Dd |||||  
Dd 2101 TTCACCTGAGGAAGGCATGGTCAACGCCCCGTTTGACTACAACCTATGACAACAGCTTCCGG 2160  
QY 6305 GTGACCAGCATGACGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTAT 6364  
Dd |||||  
Dd 2161 GTGACCAGCATGACGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTAT 2220  
QY 6365 GATGATGTGTAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGCAATTTACTATGACAT 6424  
Dd |||||  
Dd 2221 GATGATGTGTAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGCAATTTACTATGACAT 2280  
QY 6425 AACCAGATCATACACAGCTGTGATGACCCACACCAAGCAATTTGATGCATATGGCAGG 6484  
Dd |||||  
Dd 2281 AACCAGATCATACACAGCTGTGATGACCCACACCAAGCAATTTGATGCATATGGCAGG 2340  
QY 6485 ATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTAT 6544  
Dd |||||  
Dd 2341 ATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTAT 2400  
QY 6545 GATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCCAATACCACT 6604

Dd |||||  
Dd 2401 GATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACT 2460  
QY 6605 CGCTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAG 6664  
Dd |||||  
Dd 2461 CGCTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAG 2520  
QY 6665 CCACCTCTGGCGCTACAGCTACGACCTCAATGGGAAACCTGCATTAAGTGGCCTGGGAAC 6724  
Dd |||||  
Dd 2521 CCACCTCTGGCGCTACAGCTACGACCTCAATGGGAAACCTGCATTAAGTGGCCTGGGAAC 2580  
QY 6725 AGTGACGGCTCACACCACTACGGTATGACATCCCGCAGCCGATCACTCGGCTGGGTGAC 6784  
Dd |||||  
Dd 2581 AGTGACGGCTCACACCACTACGGTATGACATCCCGCAGCCGATCACTCGGCTGGGTGAC 2640  
QY 6785 GTGCAATACAAGATGGATGAGGATGGCTTCTGAGGCAGCGGGCGGTGATATCTTTGAG 6844  
Dd |||||  
Dd 2641 GTGCAATACAAGATGGATGAGGATGGCTTCTGAGGCAGCGGGCGGTGATATCTTTGAG 2700  
QY 6845 TACAACTCAGCTGGCCTGCTCATCAAGGCCCTACAACCGGGCTGGCAGCTGGAGTGTGAG 6904  
Dd |||||  
Dd 2701 TACAACTCAGCTGGCCTGCTCATCAAGGCCCTACAACCGGGCTGGCAGCTGGAGTGTGAG 2760  
QY 6905 TACCGCTACGATGGCCTGGGGCGGCGGTGTCCAGCAAGAGCAGCCACACCCACCACTG 6964  
Dd |||||  
Dd 2761 TACCGCTACGATGGCCTGGGGCGGCGGTGTCCAGCAAGAGCAGCCACACCCACCACTG 2820  
QY 6965 CAGTTCTTCTATGACACCTGACCAACCCCAACAAAGGTACCCACCTGTACAACCACTCC 7024  
Dd |||||  
Dd 2821 CAGTTCTTCTATGACACCTGACCAACCCCAACAAAGGTACCCACCTGTACAACCACTCC 2880  
QY 7025 AGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGAGCTG 7084  
Dd |||||  
Dd 2881 AGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGAGCTG 2940  
QY 7085 AGCAGTGGTGAATGATTTTACATAGCTTTGTGACAACATCGGACCCCTCTTGTCTCTTT 7144  
Dd |||||  
Dd 2941 AGCAGTGGTGAATGATTTTACATAGCTTTGTGACAACATCGGACCCCTCTTGTCTCTTT 3000  
QY 7145 AGTGAACAGGTTTGAATGATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTACATG 7204  
Dd |||||  
Dd 3001 AGTGAACAGGTTTGAATGATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTACATG 3060  
QY 7205 GATACCAACCCCACTTTTACATCATAGGCTACATAGGCTCCATGGTGGCCTCTATGATCCACTC 7264  
Dd |||||  
Dd 3061 GATACCGACCCCACTTTTACATCATAGGCTACATAGGCTCCATGGTGGCCTCTATGATCCACTC 3120  
QY 7265 ACCAAGCTTGTCCACATGGGCGGCGGAGATTATGATGTGTGGCCGACGCTGGACTAGC 7324  
Dd |||||  
Dd 3121 ACCAAGCTTGTCCACATGGGCGGCGGAGATTATGATGTGTGGCCGACGCTGGACTAGC 3180  
QY 7325 CCAGACCACGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCCTTTTAATCTCTAT 7384  
Dd |||||  
Dd 3181 CCAGACCACGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCCTTTTAATCTCTAT 3240  
QY 7385 ATGTTCAAAAACAACACCCCATCAGCAACTCCAGGACATCAAGTGTTCATGACAGAT 7444  
Dd |||||  
Dd 3241 ATGTTCAAAAACAACACCCCATCAGCAACTCCAGGACATCAAGTGTTCATGACAGAT 3300  
QY 7445 GTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTCATCCCTGGTTATCCC 7504  
Dd |||||  
Dd 3301 GTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTCATCCCTGGTTATCCC 3360  
QY 7505 AAACCAGACATGGATGCCATGGAAACCCCTCTACGAGCTCATCCACACACAGATGAAAACG 7564  
Dd |||||  
Dd 3361 AAACCAGACATGGATGCCATGGAAACCCCTCTACGAGCTTATCCACACACAGATGAAAACG 3420  
QY 7565 CAGGAGTGGGACAACAGCAAGTCTATCCCTGGGGTACAGTGTGAAGTACAGAAGCAGCTC 7624  
Dd |||||  
Dd 3421 CAGGAGTGGGACAACAGCAAGTCTATCCCTGGGGTACAGTGTGAAGTACAGAAGCAGCTC 3480  
QY 7625 AAGGCCCTTTGTCACTTAGAACGGTTTGAACAGCTCTATGGCTCCACAATCACCAGCTGC 7684

Db 3481 AAGGCCTTGTACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACAATCACCAGCTGC 3540  
Qy 7685 CAGCAGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTC 7744  
Db 3541 CAGCAGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTC 3600  
Qy 7745 AAGTTGCTTGAAGGATGGCCAGTGACCAAGAGATCATCAGTGTGGCCAAATGAGGAT 7804  
Db 3601 AAGTTGCTTGAAGGATGGCCAGTGACCAAGAGATCATCAGTGTGGCCAAATGAGGAT 3660  
Qy 7805 GGGCAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAACCTGCACTTCACC 7864  
Db 3661 GGGCAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAACCTGCACTTCACC 3720  
Qy 7865 ATTGATGGGGTGATACCCATTTGTAACCATGAAACAGGACCTTCAGAAGGTGACCTGGCC 7924  
Db 3721 ATTGATGGGGTGATACCCATTTGTAACCATGAAACAGGACCTTCAGAAGGTGACCTGGCC 3780  
Qy 7925 ATCTGGGGCTCAGTGGGGGGCGGCAACCTTGAGAAATGGGGTCAACGTCACCTGTGTC 7984  
Db 3781 ATCTGGGGCTCAGTGGGGGGCGGCAACCTTGAGAAATGGGGTCAACGTCACCTGTGTC 3840  
Qy 7985 CAGATCAACACAGTACTTAATGCGAGGACTAGACGCTACACAGACATCCAGTCCAGTAC 8044  
Db 3841 CAGATCAACACAGTACTTAATGCGAGGACTAGACGCTACACAGACATCCAGTCCAGTAC 3900  
Qy 8045 GGGGCACTGTGCTTGAACACACAGCTACGGGACAAACGTTGGATGAGGAGAGCAGCGGTC 8104  
Db 3901 GGGGCACTGTGCTTGAACACACAGCTACGGGACAAACGTTGGATGAGGAGAGCAGCGGTC 3960  
Qy 8105 CTGGAGCTGGCCCGGAGAGAGCCGTGCGCCAAAGCCGTGGGCGCGGAGCAGAGACTG 8164  
Db 3961 CTGGAGCTGGCCCGGAGAGAGCCGTGCGCCAAAGCCGTGGGCGCGGAGCAGAGACTG 4020  
Qy 8165 CGGGAAGGGAGGAAGGCGCTGCGGCGCTGGACAGAGGGGGAGAGCAGCAGGTGCTGAGC 8224  
Db 4021 CGGGAAGGGAGGAAGGCGCTGCGGCGCTGGACAGAGGGGGAGAGCAGCAGGTGCTGAGC 4080  
Qy 8225 ACAGGCGGGTGCAAGGCTACGACGGCTTTTTCGATGATCTCTGTCGAGCAGTACCCAGAA 8284  
Db 4081 ACAGGCGGGTGCAAGGCTACGACGGCTTTTTCGATGATCTCTGTCGAGCAGTACCCAGAA 4140  
Qy 8285 CTGTCAGACAGCGCCCAACATCCACTTTCATGAGACAGCGGATGGCGCGAGGTGA 8344  
Db 4141 CTGTCAGACAGCGCCCAACATCCACTTTCATGAGACAGCGGATGGCGCGAGGTGA 4200  
Qy 8345 CAGAGAGGAC 8354  
Db 4201 CAGAGAGGAC 4210

RESULT 8  
AB037723  
LOCUS AB037723 8624 bp mRNA linear PRI 10-MAY-2002  
DEFINITION Homo sapiens mRNA for KIAA1302 protein, partial cds.  
ACCESSION AB037723  
VERSION AB037723.2 GI:20521827  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.  
TITLE Prediction of the coding sequences of unidentified human genes.  
XVI. The complete sequences of 150 new cDNA clones from brain which  
code for large proteins in vitro  
JOURNAL DNA Res. 7 (1), 65-73 (2000)  
MEDLINE 20181126  
PUBMED 10718198  
REFERENCE 2 (bases 1 to 8624)  
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.  
TITLE Direct Submission

JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,  
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,  
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,  
Fax:+81-438-52-3914)  
COMMENT On May 9, 2002 this sequence version replaced gi:7242958.  
FEATURES  
Location/Qualifiers  
source  
1..8624  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="bf00104"  
/tissue\_type="brain"  
/clone\_lib="pBCSKPGW"  
/note="This sequence was replaced that of fg06864 cDNA as  
a representative cDNA sequence for KIAA1302."  
gene  
1..8624  
/gene="KIAA1302"  
CDS  
<590..3850  
/gene="KIAA1302"  
/note="Start codon is not identified."  
/codon\_start=1  
/product="KIAA1302 protein"  
/protein\_id="BAA92540.2"  
/db\_xref="GI:20521828"  
/translation="STERLDNKPETTRDYSPGRLTNTVFTPTQVSSFRSDTDSVVHVQ  
VETSSKDDVTITNLSASGAFYTLQDOVRNSYIGADGSLRLLLANGMEVALQTEPH  
LLAGTVNPTVGRNVTLPIDNGLNLEWRQKEQARGQVTFGRRLRVHNRNLLSLDF  
DRVTRTEKIYDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVYSPGGYIAGIQRGIM  
SERMEYDQAGRIPADGKTWYTYLEKSMVLLHLSQRYTFEDKNDRLSLVTMP  
NVARQTLETIRSVGYRNIYQPPENASVIQDFTEDGHLHLLTFLGTGRRVIYKYGL  
SKLAETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGLIDRQIFRTEEG  
MVNARFDYNDNSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFGKFGVIYDINQI  
ITAVMTHTKHFDAYGRMKVQYEIERSLMYWMYQYDNMGRVYVKELKVPYANTTR  
YSYEYDADGQLQTVSINDKPLWRYSYDLNGLNHLSPNSARLTPLRYDIRDITRLG  
DVQYKMDDEGLRQGGDIFEYNSAGLLIKAYNRAGSWRYRDGLGRRVSSKSSHS  
HHLOFFYADLTNPTKVTLYNHSSEISLYDLOGLHFAELSSGDEFYIACDNIGT  
PLAVFSGTGLMIKQILYTAGGEIYMDTNPNFQIIIGYHGLYDPLTKLHVMGRRDYDV  
LAGRWTSPDHELKHLSSNVMPFNLYMKNPNPISNSQDIKCFMTDVNSWLLTFGFQ  
LHNVI PGPKPDMAMEPSYELIHTQMTQEDNSKSLGVQCEVQKQLKAFVTLERF  
DQLYGSTITSQQAPKTKFASGSGVFGKGVKFKALDGRVTTDIISVANEDGRRVAAI  
LNHAHYLENHFTIDGVDTHTYFKPGSEGLAILGLSGRRRLTENGVNVTVSQINTV  
LNGRTRRYTDIQLQYCALCLNTRYGTLDEEKARVLELARQRAVRQAWAREQQRLREG  
EEGLRAWTEGEKQQLVSTGRVQGYDGFVVISVEQPELSDSANNIHFMRQSEMGRR"

ORIGIN

Query Match 44.1%; Score 3680.8; DB 9; Length 8624;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 3775; Conservative 0; Mismatches 2; Indels 83; Gaps 1;  
Qy 4578 GTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAA 4637  
Db 1 GTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAA 60  
Qy 4638 AGTTAAATACCCCATCTTCTTGGCTGTGTGTGCTGATGGGAGCTCTACGTGGCCGACC 4697  
Db 61 AGTTAAATACCCCATCTTCTTGGCTGTGTGTGCTGATGGGAGCTCTACGTGGCCGACC 120  
Qy 4698 TTGGGAACATCCGAATTCCGTTTATCCGGAAGAACAAAGCCTTTCTCAACACCCAGAAACA 4757  
Db 121 TTGGGAACATCCGAATTCCGTTTATCCGGAAGAACAAAGCCTTTCTCAACACCCAGAAACA 180  
Qy 4758 TGTATGAGCTGTCTTCACCAATTGACCAAGAGCTCTATCTGTTTATACACCCGCAAGC 4817  
Db 181 TGTATGAGCTGTCTTCACCAATTGACCAAGAGCTCTATCTGTTTATACACCCGCAAGC 240  
Qy 4818 ACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACCTACACTGGGG 4877  
Db 241 ACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACCTACACTGGGG 300  
Qy 4878 ACGGCGACATCACACTCATCACAGACAAACATGSCAACATGGTAAATGTCCGCCGAGACT 4937  
Db 301 ACGGCGACATCACACTCATCACAGACAAACATGSCAACATGGTAAATGTCCGCCGAGACT 360

QY 4938 CTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTGTACTGGGTGACCATGG 4997  
Dd |||||  
361 CTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTGTACTGGGTGACCATGG 420  
QY 4998 GCACCAACAGTGCACCTCAAGAGTGTGACCACACAAGGACACAGATGGCCATGATGACAT 5057  
Dd |||||  
421 GCACCAACAGTGCACCTCAAGAGTGTGACCACACAAGGACACAGATGGCCATGATGACAT 480  
QY 5058 ACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAACGGATGGACAACATTTT 5117  
Dd |||||  
481 ACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAACGGATGGACAACATTTT 540  
QY 5118 ATGA----- 5121  
Dd |||||  
541 ATGACAACCTCAGATGGCAGGAATTATCATCCACATTTTTTACATATGAAGTACAGAGAG 600  
QY 5122 -----GTACGACAGCTTTGGCCCGCTTGACAAATGTGAC 5154  
Dd |||||  
601 ATTAGATAACAAGCCTGAAACCAACACCGGTACGACAGCTTTGGCCCGCTGACAAATGTGAC 660  
QY 5155 CTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCA 5214  
Dd |||||  
661 CTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCA 720  
QY 5215 GGTAGAGACCTCCAGCAAGGATGATGTACCATTAACCAACCACTGTCTGCCCTCAGGCGC 5274  
Dd |||||  
721 GGTAGAGACCTCCAGCAAGGATGATGTACCATTAACCAACCACTGTCTGCCCTCAGGCGC 780  
QY 5275 CTTCTACACACTGCTGCAAGACCAAGTCCGGAAACAGTACTACATCGGGCCGATGGCTC 5334  
Dd |||||  
781 CTTCTACACACTGCTGCAAGACCAAGTCCGGAAACAGTACTACATCGGGCCGATGGCTC 840  
QY 5335 CTTGCGGCTGCTGGCCAACCGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTGCT 5394  
Dd |||||  
841 CTTGCGGCTGCTGGCCAACCGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTGCT 900  
QY 5395 GGCTGGACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACAGCTGCCCTCGACACCGG 5454  
Dd |||||  
901 GGCTGGACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACAGCTGCCCTCGACACCGG 960  
QY 5455 CCTCAACCTGGTGGAGTGGCGCCAGCGCAAAAGAGCAGGCTCGGGCCAGGTCACTGCTT 5514  
Dd |||||  
961 CCTCAACCTGGTGGAGTGGCGCCAGCGCAAAAGAGCAGGCTCGGGCCAGGTCACTGCTT 1020  
QY 5515 TGGCGCCCGCTGCGGTGCACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAAC 5574  
Dd |||||  
1021 TGGCGCCCGCTGCGGTTCACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAAC 1080  
QY 5575 ACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTGTACGACCA 5634  
Dd |||||  
1081 ACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTGTACGACCA 1140  
QY 5635 GGCGGGGGCCCGCCTCTGGTCAACCGCAGCAGGCTGAATGGTGTCAACCGTGACATA 5694  
Dd |||||  
1141 GGCGGGGGCCCGCCTCTGGTCAACCGCAGCAGGCTGAATGGTGTCAACCGTGACATA 1200  
QY 5695 CTCCTCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAATGGAATA 5754  
Dd |||||  
1201 CTCCTCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAATGGAATA 1260  
QY 5755 CGACCAAGCGGGCCCGCATCACATCAGGATCTTCGCTGATGGGAAGACATGGAGCTACAC 5814  
Dd |||||  
1261 CGACCAAGCGGGCCCGCATCACATCAGGATCTTCGCTGATGGGAAGACATGGAGCTACAC 1320  
QY 5815 ATACTTAGAAGTCCATGGTGTCTGCTACTACACAGCCAGAGGCAGTATATCTTTGAGTT 5874  
Dd |||||  
1321 ATACTTAGAAGTCCATGGTGTCTGCTACTACACAGCCAGAGGCAGTATATCTTTGAGTT 1380  
QY 5875 CGACAAGATGACCGCCTCTCTCTGTGACGATGCCCCAACCTGGCGCGGAGACTAGA 5934  
Dd |||||  
1381 CGACAAGATGACCGCCTCTCTCTGTGACGATGCCCCAACCTGGCGCGGAGACTAGA 1440

QY 5935 GACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTC 5994  
Dd |||||  
1441 GACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTC 1500  
QY 5995 AGTCATACAGGACTTCACTGAGGATGGGACCTCTTTCACACCTTTCTACCTGGGCACTGG 6054  
Dd |||||  
1501 AGTCATACAGGACTTCACTGAGGATGGGACCTCTTTCACACCTTTCTACCTGGGCACTGG 1560  
QY 6055 CCGCAGGTGATATACAAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACAC 6114  
Dd |||||  
1561 CCGCAGGTGATATACAAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACAC 1620  
QY 6115 CACCAAGGTCAGTTTCACCTATGACGAGAGCGGCAGGCATGCTGAAGACCATCAACCTACA 6174  
Dd |||||  
1621 CACCAAGGTCAGTTTCACCTATGACGAGAGCGGCAGGCATGCTGAAGACCATCAACCTACA 1680  
QY 6175 GAATGAGGGCTTCACCTGCACCATCCGCTPACCGTCAAGTTGGGCCCTGANTGACCCGACA 6234  
Dd |||||  
1681 GAATGAGGGCTTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCTGATTGACCCGACA 1740  
QY 6235 GATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCCCTTTTGTACTACAACTATGACAA 6294  
Dd |||||  
1741 GATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCCCTTTTGTACTACAACTATGACAA 1800  
QY 6295 CAGCTTCCGGGTGACAGCATGCAAGCTGTGATCAACGAGACCCCACTGCCATTGATCT 6354  
Dd |||||  
1801 CAGCTTCCGGGTGACAGCATGCAAGCTGTGATCAACGAGACCCCACTGCCATTGATCT 1860  
QY 6355 CTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGATTGGGAGTGTGGTGTCAATTA 6414  
Dd |||||  
1861 CTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGATTGGGAGTGTGGTGTCAATTA 1920  
QY 6415 CTATGACATTTAACAGATCATCACCACAGCTGTCTATGACCCACACCAAGCATTTTGATGC 6474  
Dd |||||  
1921 CTATGACATTTAACAGATCATCACCACAGCTGTCTATGACCCACACCAAGCATTTTGATGC 1980  
QY 6475 ATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGAC 6534  
Dd |||||  
1981 ATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGAC 2040  
QY 6535 CGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGGAGCTGAAGTAGGACCCCTACGC 6594  
Dd |||||  
2041 CGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGGAGCTGAAGTAGGACCCCTACGC 2100  
QY 6595 CAATACCACCTCGTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCAT 6654  
Dd |||||  
2101 CAATACCACCTCGTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCAT 2160  
QY 6655 CAATGACAAAGCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAG 6714  
Dd |||||  
2161 CAATGACAAAGCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAG 2220  
QY 6715 CCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGGACCGCATCACTCG 6774  
Dd |||||  
2221 CCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGGACCGCATCACTCG 2280  
QY 6775 GCTGGTGACGTGCAATACAAGATGGATGGATGGCTTCCTGAGSCAGCGGGCGGTGA 6834  
Dd |||||  
2281 GCTGGTGACGTGCAATACAAGATGGATGGATGGCTTCCTGAGSCAGCGGGCGGTGA 2340  
QY 6835 TATCTTTGAGTACAACTCAGCTGGCCCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTG 6894  
Dd |||||  
2341 TATCTTTGAGTACAACTCAGCTGGCCCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTG 2400  
QY 6895 GAGTGTGAGTACCGCTACGATGGCCCTGGGGCGGCGCTGTCCAGCAAGAGCAGCCACAG 6954  
Dd |||||  
2401 GAGTGTGAGTACCGCTACGATGGCCCTGGGGCGGCGCTGTCCAGCAAGAGCAGCCACAG 2460  
QY 6955 CCACCACCTGCAGTTCTTCTATGACAGCTGACCAACCCCAAGAGTCAACCCACCTGTA 7014  
Dd |||||  
2461 CCACCACCTGCAGTTCTTCTATGACAGCTGACCAACCCCAAGAGTCAACCCACCTGTA 2520  
QY 7015 CAACCACCTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGAAGGACACCTCTTTGC 7074

Db 2521 CAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGC 2580  
Qy 7075 CATGGAGCTGAGCAGTGGTGATGAGTTTACATAGCTTGTGACAAACATCGGGACCCCTCT 7134  
Db 2581 CATGGAGCTGAGCAGTGGTGATGAGTTTACATAGCTTGTGACAAACATCGGGACCCCTCT 2640  
Qy 7135 TGCTGTCTTTAGTGGAAACAGTTTGTATGATCAAGCAAAATCCTGTACACAGCCTATGGGA 7194  
Db 2641 TGCTGTCTTTAGTGGAAACAGTTTGTATGATCAAGCAAAATCCTGTACACAGCCTATGGGA 2700  
Qy 7195 GATCTACATGGATACCAACCCCAACTTTTCAGATCATATAGGCTACCATGGTGGCCTCTA 7254  
Db 2701 GATCTACATGGATACCAACCCCAACTTTTCAGATCATATAGGCTACCATGGTGGCCTCTA 2760  
Qy 7255 TCATCCACTCACCAAGCTTGTCTCACATGGCCGGCGAGATTATGATGTGTGGCCGGACG 7314  
Db 2761 TGATCCACTCACCAAGCTTGTCTCACATGGCCGGCGAGATTATGATGTGTGGCCGGACG 2820  
Qy 7315 CTGGACTAGCCCCAGACCCAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCTTT 7374  
Db 2821 CTGGACTAGCCCCAGACCCAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCTTT 2880  
Qy 7375 TAATCTCTATATGTTCAAAAACAACAAACCCCATCAGCAACTCCCAGGACATCAAGTGTCT 7434  
Db 2881 TAATCTCTATATGTTCAAAAACAACAAACCCCATCAGCAACTCCCAGGACATCAAGTGTCT 2940  
Qy 7435 CATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTCATCCC 7494  
Db 2941 CATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTCATCCC 3000  
Qy 7495 TGGTTATCCCAAAACCAGACATGGATGCCATGGBAACCCCTCCTACGAGCTCATCCACACACA 7554  
Db 3001 TGGTTATCCCAAAACCAGACATGGATGCCATGGBAACCCCTCCTACGAGCTTATCCACACACA 3060  
Qy 7555 GATGAAAACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACA 7614  
Db 3061 GATGAAAACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACA 3120  
Qy 7615 GAAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACAAT 7674  
Db 3121 GAAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACAAT 3180  
Qy 7675 CACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTTGG 7734  
Db 3181 CACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTTGG 3240  
Qy 7735 CAAGGGGGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTGGC 7794  
Db 3241 CAAGGGGGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTGGC 3300  
Qy 7795 CAATGAGGATGGCGGAAGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAACCT 7854  
Db 3301 CAATGAGGATGGCGGAAGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAACCT 3360  
Qy 7855 GCACTTCACCATTGATGGGCTGATACCCATTACTTTGTGAAACCCAGGACCTTCAGAAGG 7914  
Db 3361 GCACTTCACCATTGATGGGCTGATACCCATTACTTTGTGAAACCCAGGACCTTCAGAAGG 3420  
Qy 7915 TGACCTGGCCATCCTGGGCTCAGTGGGGGGCGGGAACCCCTGGAGATGGGCTCAACGT 7974  
Db 3421 TGACCTGGCCATCCTGGGCTCAGTGGGGGGCGGGAACCCCTGGAGATGGGCTCAACGT 3480  
Qy 7975 CACTGTGTCCCAGATCAACACAGTACTTAATGSCAGGACTAGACGCTACACAGACATCCA 8034  
Db 3481 CACTGTGTCCCAGATCAACACAGTACTTAATGSCAGGACTAGACGCTACACAGACATCCA 3540  
Qy 8035 GCTCCAGTACGGGGCACTGTGCTTGAACACACGCTACGGGACAACCGTTGGATGAGGAGAA 8094  
Db 3541 GCTCCAGTACGGGGCACTGTGCTTGAACACACGCTACGGGACAACCGTTGGATGAGGAGAA 3600  
Qy 8095 GGCACGGGTCTTGGAGCTTGGCCCGGCAGAGAGCCGTGCGCCCAAGCGTGGGCCCGCGAGCA 8154

Db 3601 GGCACGGGTCTTGAGTGGCCCGGCAGAGAGCGCTGCGCCAAGCGTGGGCCCGCAGCA 3660  
Qy 8155 GCAGAGACTGCGGGAAGGGGAGGAAGCCCTGCGGGCCTGGACAGAGGGGGAAGAGCAGCA 8214  
Db 3661 GCAGAGACTGCGGGAAGGGGAGGAAGCCCTGCGGGCCTGGACAGAGGGGGAAGAGCAGCA 3720  
Qy 8215 GGTGCTGAGCACAGGCGGGTCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCA 8274  
Db 3721 GGTGCTGAGCACAGGCGGGTCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCA 3780  
Qy 8275 GTACCCAGAACTGTACAGACAGCGCCAAACAACATCCACTTCATGAGACAGAGCGAGATGG 8334  
Db 3781 GTACCCAGAACTGTACAGACAGCGCCAAACAACATCCACTTCATGAGACAGAGCGAGATGG 3840  
Qy 8335 CCGGAGGTGACAGAGAGGAC 8354  
Db 3841 CCGGAGGTGACAGAGAGGAC 3860

RESULT 9  
AB025412  
LOCUS Mus musculus mRNA for Ten-m3, complete cds. 8964 bp mRNA linear ROD 08-MAY-1999  
DEFINITION AB025412  
ACCESSION AB025412  
VERSION AB025412.1 GI:4760779  
KEYWORDS Ten-m3.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (sites)  
Ohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,  
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.  
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane  
proteins expressed in many tissues  
J. Cell Biol. (1999) In press  
REFERENCE 2 (bases 1 to 8964)  
Ohashi,T.  
AUTHORS Direct Submission  
TITLE Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University  
JOURNAL Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,  
Okayama, Okayama 700-8252, Japan  
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128,  
Fax:+81-86-222-7768)

FEATURES  
source Location/Qualifiers  
1..8964  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="Balb/c"  
/db\_xref="taxon:10090"  
/tissue\_type="brain"  
/dev\_stage="adult"  
1..8964  
/gene="Ten-m3"  
50..8197  
/gene="Ten-m3"  
/codon\_start=1  
/product="Ten-m3"  
/protein\_id="BAA77398.1"  
/db\_xref="GI:4760780"  
/translation="MDVKRRPYCSLTKSRRKERYTNSADNEECRVPTQKSYSS  
ETLKFHDYSRLLYGNRVKDLVHREADEYTRQGNFTLQLGVCESATRRGVAFCAE  
MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRGVKSGRSSCLSSRSNSALTLD  
EHENRSDSESEQPSNNPGQTLPLPSSHQHPAQHHPSTSLNRSLSLNNRNSQSPAP  
PAALPAELQTTPEVQLQDSWVLGNSVPLESRHFLFKTGTCTPLFSTATPGYTMASG  
SVYSPTRPLPRNTLSRAFAFKFKSSKYCSWRCTALCAVGSVLLAILLSYFIAMHLF  
GLNWLHQQTENDTFENGKNSDTPNTVSLPSGDNKLGGFTHENNTIDSGELDIGR  
RAIQEVPGIFWRSOLFIDQPOFLKFNISLOKDALIGVYGRKGLPPSHTQYDFVELLD  
GSRLLAREQNLVSEERAGQARSVSLHEAGFIQYLDSGIWHLAFYNDKGNPEQVSFN  
TIVIESVVECPRNCHNGECVSGTCHCFPGFLGPDCSRACPVLCSGNGQYSGKRCCLC  
FSGWKGTEDVPTTQCIDPQCGGIGICIMGSCACNSGYKGENCEBADCLDPGCSNHGV  
CIHGECHNPGWGSNCIELKTMCADQCSGHGTYLQESGSCDTPNWTGPDCSNEICS  
VDCGSHGVCMGGSCRCCEGTGTGACNQRACHPRCAEHGTCKDKKCECSQGWNGEHTCI

gene  
CDS

AHYLDKI VKEGCPGLCNSNGRCTLDQNGWHVCVQPGWRGAGCDVAMETLCTDTSKDNEG  
 DGLIDCMPDPCCLQSSCQNPYCRGLPDPQDIISQSLQTPSQQAASFPYDRISFLIGS  
 DTHVLPGESFPFNKSLASVIRGVLQTPADTPLIGNVSVFLHYSEYGYITRQDGMFDL  
 VANGASLTLVFPERSPELTQYHTVWIPKNVYVYMDTLVMKKEGTLPSCDLSGFVRPS  
 PII VSSPLSTFFRSPSPEDSPII PETQVLHEETITPGTDLKLSYLSRGAAGKSVLKIT  
 MTQAVIPFNLKMHVMAVGRLFQKWFPA SPNLAYTFIWDKTDAYNQKVYGLSEAV  
 SVGYEYESCLDLTLWEKRTAVLOQYELDASNMGGWTLDKHHVLDVONGILYKNGENQ  
 FVSEQPPVSSIMNGNRRRSISCPSCNQAGDNKLAPALACDGLSIVYDFKNLYR  
 RIFPSQPNVSVLELRNCKFRHSNPAHRYLADPTVTDGLYVSDTNRIRYRPKSLTGV  
 AKDLTKNAEVVAGTGEQCLPFDEARCDGGKAVEATL MSPKGMADKNGLIYFVDDGTW  
 IRKVDONGIISTLLGNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDN  
 NVLQITENRQVITAAAGRPMHCQVGPVEYVPGKHAVQTTLSEATLAVSYSGVLYITE  
 TDEKKNIRQVTTDGEISLVLAGIPSECDKNDANCDCYQSGYAKDAKLNAPSSLA  
 ASPDGTLYIADLGNIRIRAVSKNKPILNLSMNFYEVASPTDQELYTFDINGTHQYTVSL  
 VTGDLYNFPYSYNDNDVTA VTDNSGNTLIRIRDPNRMPPVRVSPNQVILWLTIGTNGC  
 LKSMTAQGLELVLFTYHGNSGLLAKTSDETGTTFFDYDSEGRNLNYTPTGVVTLNLH  
 GMDKAITVDLIESSREEDVSI TSNLSSIDSFPYTMQDQLRNSYQIGYDGLSRI FYAS  
 GLDSHYQTEPHVLAGTANPTAKENNTLPGENGONLVEWRFRKEAQGVNFGFKLR  
 VNGRNLLSVD FDR TT KTEKIYDDHRKFLRLRIAYDTSCHPTLWLPSSKLMANVTYSST  
 GQIASIQRTTSSEKVDYDSQGRIVSRVPADGKTWSYTYLEKSMVULLHSQRQYIFEYD  
 MWDLASAITMPSVARHTMQTIRSGYNYNIYNPESNASIITDNEEGELTLQTAFLGT  
 SRRVLFAKYRRQTRLSEILYDTSVSTYDNETAGVLKTVNLQSDGFTLIRYRQIGPLI  
 DRQIFPSEDDGMVNARFYSYDNSFRVTSMQGVINETPLPDLQFDDISGKVEQFGK  
 FGVLYYDINQIISTAVMTYTKHFDAHGRIKEIQEYIFRSLMYWITI QYDNMGRVTKRE  
 IKICFPANTTKYAYEYDVGQQLQTVLNEKIMWRNYDNLNGNHLNPPSSARLTPLR  
 YDLDRITRLGDVQDGLFRLRGRTGEIFEYSSKGLLTRVYSKSGSWTVIYRYDGL  
 GRRVSSKTSLGQHLQFFYADLTYPTRITHVYNHSSSEITSLYDLQGLHFAMEISSGD  
 EFYIASDNTGTPLAVFSSNGLMKQOTYATGEIYFDSNVDFQLVGFHGGLYDPLTK  
 LTHFERDYDILAGRWTTDPIETWKRIGKDPAPFNLYMFRNNNPSAKIHDDKDYITDV  
 NSWLVTGFGHLHNAIPGFFVPVKFDLT EPSYELVKSQWEDVPPFVGQQQVARQAKAF  
 LLSLGMAEVQVSRRKAGAEQSWLWFAVKSILGKGVMLAVSPGQVQTNVLNIANEDCI  
 KYAAVLNNAPYLENLHFTIEGKDTHYFIKTTTPESDLGTLRLTSRKALENGINVTVS  
 QSTTVNGRTRRRFADVEMQFGALALHVRVGMTLDEEKARILEQARQALARAWAREQQ  
 VRVDGEEGARLWTEGEKRQLLSAGKVQGYDGYVLSVEQYPELADSA NNIIQLRQSEI  
 GRV "

## ORIGIN

Query Match 42.4%; Score 3544.6; DB 10; Length 8964;  
Best Local Similarity 65.6%; Pred. No. 0;  
Matches 5458; Conservative 0; Mismatches 2679; Indels 186; Gaps 12;

QY	35	ATGGACGTGAAGGAGAGGAAGCCCTTACCGCTCGCTGACCCGGCCG---	CGCAGCCGAG	91
Db	50	ATGGATGTGAAGGAACGCAGGCCCTTACTGCTCTTGACCAAGAGCAGACGGGAAAGGAA	109	
QY	92	CGCCGTACACACAGCTCGTCCGGGACAGCGAGGAGGCAAGGCC---	CGCAGAAATCG	148
Db	110	AGCGCTATACAAAATTCTGCCGGCAATGAGGAGTGTAGGTTCCCCACGCAGAAAGTCC	169	
QY	149	TACAGCTCCAGCGAGACCCCTGAAGSCCTACGACCAGG---	ACGCCGCCCTAGCCTATGGC	205
Db	170	TATAGTTCAGTGAAACCTTGAAGCTTCGATCATGATTATTCACGGCTGCTTTATGGA	229	
QY	206	AGCCGCTCAAGGACATTGTGCCGAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAAC	265	
Db	230	AACAGAGTAAAGGATTTGGTCCACAGAGAAGCCGACGAGTATACTAGACAAGGACAGAA	289	
QY	266	TTTACCCCTGCGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGG	325	
Db	290	TTTACCCCTAAGGCAGTTAGGAGTGTGTGAATCCGCAACTCGAAGAGGAGTGGCAATCTGT	349	
QY	326	ACAGACATTGGCCTGCCCAATGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATG	385	
Db	350	GCGGAATGGGGCTCCCTCACAGAGGTTACTCCATCAGTGCAGGGTCAGATCGGGATACG	409	
QY	386	GAGGTGACACGGTGTCTCCCTGAGCAACCCCGTGGTCTGTGGGGCCGGAGCACACGG	445	
Db	410	GAAACGAAGCAGTGATGTCCCTGAGCATGCCATGAGACTTTGGGGCAGGGGGGTCAAA	469	
QY	446	TCAGGGCGCAGTCTCTGCTGTCCAGCCGGGCCAAATTCCAATCTCACACTCACCCGACACC	505	
Db	470	TCGGGCGCAGTTCCTGCTGTCAAGCCGGTCCAACTCCGCCCTCACCCCTGACAGACACG	529	

Db 1457 CGCAGGCGAGATCTGTGAGCCCTGCAGGAGTGGCTTCATCCAGTACTGGATCTGGA 1516  
QY 1646 ATCTGGCACTTGGCTTTTACAATGACGGAAAGGAGTCAAGAGTGGTTCTCTCACC 1705  
Db 1517 ATCTGGCACTTGGCTTTTATAACGACGGGAAACCCAGAGCAGGTCTCTTTAACACG 1576  
QY 1706 ACTGCCATTGAGTCGGTGGATAACTGCCCCAGCACTGCTATGGCAATGCTGACTGCATC 1765  
Db 1577 ATCGTTATAGAGTCTGTGGTGAATGCCCCCGGAATGCCATGGAATGGAGAGTGTGT 1636  
QY 1766 TCTGGGACCTGCCACTGCTTCTTGGGTTTCTCTGGGCCCCGACTGTGGCAGAGCCTCCTGC 1825  
Db 1637 TCTGGAACCTTGCCATTGTTTCCCCGGGTTTCTAGGTCGGATTGTTCAAGAGCAGCCTGT 1696  
QY 1826 CCCGTGCTCTGTAGCGGAATGGCCCAATACATGAAGAGCAGATGCTTGTGCCACAGTGGC 1885  
Db 1697 CCGTGTCTGTAGTGGCAACGGGCAATATCTCAAGGGCCGCTGCCCTGTGCTTCAGTGGC 1756  
QY 1886 TGGAAAGGCGCTGAGTGGATGTGCCCAACCAACCACTGATGATCGATGTGGCCTGCAGCAAC 1945  
Db 1757 TGGAAAGGCAACCGAGTGTGACGTGCCGACGACCCAGTGCAATTGACCCGCGAGTGGCGGGT 1816  
QY 1946 CATGGCACCTGCATCACGGCACCTGCATCTGCAACCTTGGCTACAAGGCGGAGAGCTGT 2005  
Db 1817 CGTGGGATTGTCATCATGGGCTCTTGGCTTGTAACTCGGGATACAAAGAGAAACTGT 1876  
QY 2006 GAGGAAGTGGACTGCATGGACCCCAACATGTTACGGCCGGGTGTCTGCTGAGAGGCGAA 2065  
Db 1877 GAGGAAGCGGACTGTCTAGACCCCTGGATGTTCTAATCACGGGTGTGTATCCATGGGAA 1936  
QY 2066 TGCCATTGCTTTGTGGATGGGAGGCAACCACTGGCAGACCCCGGCTTTCAGCTGTGACCCA 2125  
Db 1937 TGTCACTGCAATCCAGGCTGGGTGGCAGCACTGTGAATACTGAAGACTATGTGTGCA 1996  
QY 2126 GACCACTGTTTCAAGCCACGGAACTTCTCCCGGACACCCGGCTTTCAGCTGTGACCCA 2185  
Db 1997 GACCACTGCTCAGGCCACGGGACTTACCTTCAAGAAAGCGGCTCCTGCACCTTGGACCCA 2056  
QY 2186 AGCTGGACTGGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGC 2245  
Db 2057 AATTGGACTGGCCCCGACTGTCAATGAATATGTTCACTGGACTGCGGCTCACACGGC 2116  
QY 2246 GTGTGCTAGGGGGCACCTGCGCTCGAGGATGGCTGGATGGGGGCGAGCCCTGCGACCCAG 2305  
Db 2117 GTCTGCATGGGGGCTCCTGTGCTGTGAAGAAGGTGGACCGGCCCGCGGTGAATCAG 2176  
QY 2306 CGGGCTGCCACCCGCGCTGTCCCGAGCATGGGACCTGCCGCGACCGCAAGTGGAGTGC 2365  
Db 2177 AGAGCTTGCCACCTCGCTGTGAGCAGCGGACCGGACGTGCAAGGACGGCAAGTGGAGTGC 2236  
QY 2366 AGCCCTGGCTGGAATGGCGGAACACTGCACCATCGCTCACTATCTGGATAGGGTAGTTAA 2425  
Db 2237 AGCCAAAGGATGGAACGGAGAGCACTGCACAAATTGCTCACTATTGGATAAGATAGTTAA 2296  
QY 2426 GAGGTTGCCCTGGTGTGCAATGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGG 2485  
Db 2297 GAGGTTGCCCCCGGCTTGTGCAACAGCAATGGGAGATGCACACTGGACCCAAACCGGCTGG 2356  
QY 2486 CACTGGCTGTCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCC 2545  
Db 2357 CACTGGCTTTCAGCCAGGCTGGAGAGGAGCAGGCTGTGACGTAGCCATGGAGACCCCTC 2416  
QY 2546 TGGGTTGACAGCAAAAGCAATGATGGAGATGGCCCTGGTGGACTGCATGGACCCCTGACTGC 2605  
Db 2417 TGTACAGACAGCAAAAGCAACGAAGGAGACGGAATGACTGATGCTGATGGATCCTGATTGC 2476  
QY 2606 TGCCTCCAGCCCTGTGCTGCAATCAACCGCTGTGCTTGGCTCCCTAACCCCTCTGGAC 2665  
Db 2477 TGCCTCCAGAGCTCCTGCCAAACACAGCCCTACTGTGCTGGCTTGCCTGATCCTCAGGAT 2536  
QY 2666 ATCATCAGGAGACACAGGTCCCTGTGTACAGCAGAACCTACACTCCTCTTATGACCGC 2725

Db 2537 ATCATTAGCCAAAGCCTTCAGACACCACTCTCAGCAAGCTGCCAAGTCCCTTCTATGACCGA 2596  
QY 2726 ATCAAGTTCTCTGTCGGCAGGGACAGACGACGACATAATCCCGGGGAGAACCCCTTTGAT 2785  
Db 2597 ATCAGTTTCTGATGATGATCGATAGCACCCACGTCCTCCCTGGAGAAAGTCCGTTCAAT 2656  
QY 2786 GGAGGGCATGCTTGTGTTATTCGTGGCCAAAGTATGATCATCAGATGGAACCCCTGGTT 2845  
Db 2657 AAGAGTCTTGGCTCCGTCATCAGAGGCCCAAGTACTAACAGCTGATGGAACCCCACTTAT 2716  
QY 2846 CGTGTGAACATCAGTTTGTCAATAACCTCTCTTTGGATATACATCAGCAGGCAAGAT 2905  
Db 2717 GCGTCAACGTGTGCTTTTACACTACTCGGAATATGATATACCATACCCGCCAGGAT 2776  
QY 2906 GGCAGCTTTGACTTGTGACAAATGGCGGCATCTCCATCATCTCGGCTTCGAGCGGCA 2965  
Db 2777 GGAATGTTTGAATGCTGGCAATGGTGGCGCTTCTCTGACTTTGGTATTTGAGCGTTC 2836  
QY 2966 CTTTTCATCACACAGAGACACACCTGTGGCTGCCATGGGATCGCTTCTTTGTCTATGGAA 3025  
Db 2837 CCATTCCTCACTCAGTACCACTGTGTGGAATCCCTGGAATGTCTTTATGTATGATGGAT 2896  
QY 3026 ACCATCATCATGAGATGAGAGGAAATGAGATTCCTCCAGCTGTGACCTGAGCAATTTTGC 3085  
Db 2897 ACCCTGTCTATGAAGAAAGAGGAGAACGACATTCCTCCAGCTGTGACCTCAGTGGCTTTGTG 2956  
QY 3086 CGCCCCAACCCAGTGTCTCTCCATCCCACTGACAGTCTTTCGCCAGCTCTGTGACAG 3145  
Db 2957 AGGCCAAGTCCCATCATGTGTCTTTCACCGTTATCCACCTTCTTCCAGTCTTCCCTGAG 3016  
QY 3146 AAAGCCCCCATTTGTCGGAAATTCAGGCTTTTGACGAGGAAATCTCTATCTCTGGCTGC 3205  
Db 3017 GACAGCCCCATCATCCCGAGACACAGGTCTCTGATGAAGAAACCAATTCAGGAACA 3076  
QY 3206 AAGATGAGGCTGAGTACTGAGCAGCCGACCCCTGGCTACAAATCTGCTCCTGAGGATC 3265  
Db 3077 GATTGAAACTTTCCTACTGAGTTCAGAGCGGAGGTTACAAGTCAAGTCTTAAGATT 3136  
QY 3266 AGCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGCTAGCGGTG 3325  
Db 3137 ACCATGACCCAGGCGCTCATACCGTTTAACTCATGAAGTCCATCTGATGCTGCGCGTG 3196  
QY 3326 GAGGCGCGCTCTTCAAGAAAGTGTTCGCTGCAGCCCCAGACCTGTCTCTATTTTCAAT 3385  
Db 3197 GTTGGAGACTCTTCCAGAAAGTGGTTTCTGCTCCGCAAACTTGGCCTACAGCTTCATC 3256  
QY 3386 TGGACAAGACAGACGCTCTAACACCAAGAGGTGTGGGCTTTTTCAGAAAGCTTGTTC 3445  
Db 3257 TGGATAAGACGGACGGCATATAATCAGAAAGTCTACGGCTTGTGAGAGGAGTGTGTCC 3316  
QY 3446 GTGGTTATGAATATGAATCTGCCCCAGATCTAATCTGTGGGAAAAAGAACACAGTG 3505  
Db 3317 GTCGATAACGATACGAGTCTGTGGACTTGGACCTGACTCTCTGGGAAAGAGGAGTCCGTT 3376  
QY 3506 CTGAGGGCTATGAAATGACCGCTCCAAGCTTGAGGATGGAGCCTAGACAAACATCAT 3565  
Db 3377 TTGCAAGGCTATGAGTTGGATGCTTCGAACATGGCGGCTGGACGTTGGACAAGCACCAT 3436  
QY 3566 GCGCTCAACATTCAAAGTGGTATCTCTGCACAAAGGAATGGGGAGAACGATTTGTGTCT 3625  
Db 3437 GTACTGGACGTTTCAAGACGCTATATAACAAAGGAATGGAGAAATCAGTTTCTCTCT 3496  
QY 3626 CAGCAGCTCCTGTCTATGGAGCATCATGGCAATGGGCGCCGAGAGATCTCTCTGC 3685  
Db 3497 CAGCAGCTCCGTTGGTTCAGCAGCATCATGGTAAATGGTGGAGCGTATGATCTCATGC 3556  
QY 3686 CCCAGTCAACGGCTTGTGACGGCAACAAAGCTCTGGCCCCAGTGGCCCCACCTGT 3745  
Db 3557 CCAAGTTGCAATGGTCAAGCTGACGGGAACAACTCTGGCACCCGTCGCGCTTGCCTGT 3616  
QY 3746 GGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTCCCTCT 3805  
Db 3617 GGGATCGACGGCAGTCTATACGTAGGGGATTTCAATTACGTCCGGCGGATATTCCTCT 3676



Db 5834 AACATCTACAATCCCCAGAAAGCAATGCCCTCTATCATCACCGACTACAACAGGAAGGG 5893  
QY 6023 CACCTCCTTACACACCTTCTACCTGGGCACTGGCCGAGGGTGATATACAAGTATGGCAA 6082  
Db 5894 CTGCTCTGCAACACAGCTTTCCTGGGAACGAGTCGGAGGGTCTTATTCAAGTATAGAAG 5953  
QY 6083 CTGTCAAAGCTGGCAGAGACGCTCTATGACACCACCAAGGTCAAGTTTCACCTATGACGAG 6142  
Db 5954 CAGACCAGGCTATCAGAAATTTATACGACAGCACACAGAGTCAGTTTACCTACGACGAA 6013  
QY 6143 ACGCAGGCGATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACCTGCACCATCCGC 6202  
Db 6014 ACAGCGGAGTCCTGAAAACAGTAACCTTCAGAGTGATGGTTTTATTGACCATTAGA 6073  
QY 6203 TACCGTCAGATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACCTGAGGAAGCATG 6262  
Db 6074 TACAGGCAAAATGGTCCCTGATTGACAGACAGATTTCCGCTTCAGCGAGGATGGAATG 6133  
QY 6263 GTCAACGCCCTTTTGACTACAACTATGACAAACAGCTTCCGGGTGACCAAGCATGCAGGCT 6322  
Db 6134 GTAAATGCGAGATTGACTATAGCTACGACAAACAGCTTTCGAGTGACCAAGCATGCAGGT 6193  
QY 6323 GTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTCAAGCAAG 6382  
Db 6194 GTCATCAATGAACACCACTGCCCCATTGATCTATACAGTTTGTATGACATCTCTGGCAA 6253  
QY 6383 ACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACAGATCATCACCACA 6442  
Db 6254 GTCGAGCAGTTTGGAAAATTCGGAGTGATATACTACGACATCAACCAAAATCATTTCCACG 6313  
QY 6443 GCTGTCAATGACCCACACCAAGCATTTTGATGCTATGGCAGGATGAAGGAGTGCAGTAT 6502  
Db 6314 GCCGTGATGACTTATACAAAGCACTTTGATGCTCATGGGCGCATCAAGGAGATCCAATAT 6373  
QY 6503 GAGATCTTCGCTCGCTCATGTAAGTAGGACCCCTACGCCAATACCACCTCGCTACTCTATGAGTAT 6622  
Db 6434 ACCAAGAGAGAGATTAAATTTGGGCCCTTTTGCCAACTACCAAAATACGCGTACGAGTAC 6493  
QY 6623 GATGCTGACGGCCAGCTGACAGACAGTCTCCATCAATGACAAAGCCACTCTGGCGCTACAGC 6682  
Db 6494 GACGTGATGGACAGCTCCAAACAGTTTACCTAAACGAAAAGATCATGTGGCGGTACAAC 6553  
QY 6683 TACGACCTCAATGGGAACCTGCATTTACTGAGCCCTGGGAACAGTGCAAGCTCACACCA 6742  
Db 6554 TACGACCTAAATGGAAACCTCCACTTGTCTCAACCCGACGACGCGCCGCTGACCCCT 6613  
QY 6743 CTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGAT 6802  
Db 6614 CTGCGCTATGACCTGCGGACAGAAATCAACCGCTGGGCGATGTTTCAGTACCGCTGGAT 6673  
QY 6803 GAGGATGGCTTCTGAGGACGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTG 6862  
Db 6674 GAAGATGGTTTCTGCTGCTCAGAGGGGCACTGAAATTTTGAATACAGCTCCAAAGGGCTT 6733  
QY 6863 CTCAATCAAGGCCCTACAAACCGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCTG 6922  
Db 6734 CTGACTCGAGTCTACAGTAAAGCAGTGGCTGGACAGTGATCTATCGGTACGACGGCCTG 6793  
QY 6923 GGGCGCGCTGTCCAGCAAGACGACGCCACAGCCACCACTGCAAGTTCTTCTATGACAGAC 6982  
Db 6794 GGAAGACGTGTTTCTAGCAAAACCAAGCTTGGACAGACCTTCAGTTTCTACGCCGAC 6853  
QY 6983 CTGACCAACCCCAAGGTCAACCACTGTACAAACCACTCCAGCTCTGAGATCACCTCC 7042  
Db 6854 CTGACATACCCCAAGAAATTAATCAGCTCTACAACTTCCAGTTTCCAGTTCAGAAATCACCTCC 6913  
QY 7043 CTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTATGAGTTT 7102

Db 6914 CTGTACTATGACCTCCAAGGACATCTCTTCGCCATGGAGATCAGCAGTGGGATGAGTTC 6973  
QY 7103 TACATAGCTTGTGACAAACATCGGGACCCCTCTTGTGTCTTTTAGTGGAAACAGGTTTGATG 7162  
Db 6974 TACATCGCCTCGGACAAACACGGGGACACCGCTGGTGTTTTTTCAGCAGCAACGGGCTCATG 7033  
QY 7163 ATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTACATGGATACCAACCCCAACTTT 7222  
Db 7034 CTGAAAACAGACCCAGTACATGCTTATGGTGAGATCTACTTTGACTCCAAACGTCGACTTT 7093  
QY 7223 CAGATCATATAGGCTACCATGCTGGCCCTCTATGATCCACTCACCAAGCTTGTCCACATG 7282  
Db 7094 CAGCTGGTAATTGGATTCCACGGGGGCTTGTATGACCCGCTCACCAAACTAATCCACTTT 7153  
QY 7283 GGCAGCGGAGATTATGATGTGTGCTGGCCGACGCTGGAATAGCCCAAGACACGAGCTGTGG 7342  
Db 7154 GGAGAAAAGAGATTATGACATTTTGGCGGAAGATGGAACACACCGGACATGAAATCTGG 7213  
QY 7343 AAGACCTTAGTAGCAGCAACGTCATGCCCTTTTAATCTCTATATGTTCAAAAACAAAC 7402  
Db 7214 A---AAAGGATCGGAAAGGACCTCTCTCTTTAAACCTGTATATGTTTCGGAATAACAAC 7270  
QY 7403 CCATCAGCAACTCCAGGACATCAAGTGCTTCAAGACAGATGTTAAACAGTGGTGTCTC 7462  
Db 7271 CCGCGAGCAAAATCCATGATGTGAAAGATTACATCAGGATGTTAAACAGTGGTGTGGTG 7330  
QY 7463 ACCTTTGGATTCCAGCTACACAAACGTCATCCACACAGATGAAAACGAGGATGGGACAAAC 7522  
Db 7331 ACGTTTGGCTTCCATCTGCACAAATGCTATTCTCTGGATTCCCTGTTCCCAAAATTTGATTTA 7390  
QY 7523 ATGGAACCTCTCTACGAGCTCATCCACACAGATGAAAACGAGGATGGGACAAACAGC 7582  
Db 7391 ACTGAGCCTTCTTATGAGT-----TGTAAGAGTCAACAGTGGGAAGATGTG 7438  
QY 7583 AAGTCTATCTCGGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCCCTTTGTACCTTA 7642  
Db 7439 CCGCCCATCTTTGGAGTTCAGCAGCAAGTGGCAAGGCAAGCCAAAGGCCCTTCTGTCCCTG 7498  
QY 7643 GAACGGTTTGACCGCTCTATGGCTCCACAATCACCAAGCTGCCAGCAGGCTCCAAAGACC 7702  
Db 7499 GGAAGATGGCCGAGGTGCAGGTGAGCCGACGCAAGAGCTGGCGCCGAG---CAGTCGTGG 7555  
QY 7703 AAGAAGTTTGCATCCAGCGGCTCAGTCTTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGGAT 7762  
Db 7556 CTGTGGTTCCCAACGGTCAAGTCGCTCATCGGCAAGGGCGTCTATGCTGGCCGTGAGCCAA 7615  
QY 7763 GCGCAGTGACCAACAGACATCATCAGTGTGGCCCAATGAGGATGGGCGAAGGTTGCTGCC 7822  
Db 7616 GCGCGCTGCAGACCAACGCTGCTCAACATCGCCAAACGAGGACTGCATCAAGGTGGCGCG 7675  
QY 7823 ATCTGAACCATGCCCACTACCTAGAGAACCTGCACTTCAACATTTGATGGGTGGATACC 7882  
Db 7676 GTGTCAACAAACGCCCTTCTACCTGGAGAACCTGCACTTCAACATCGAGGGCAAGGACACA 7735  
QY 7883 CATTAATTTGTGAAACCAAGACCTTCAAGAGGTGACCTGGCCCATCTCTGGCCCTCAGTGGG 7942  
Db 7736 CACTACTTCATCAAGACCAACACCCGAGAGCGGACCTGGGCACTGCGGCTGACGAGC 7795  
QY 7943 GGGCGCGAACCCCTGGAGAAATGGGTCAACGTCATCTGTGTCCCAGATCAACACAGTACTT 8002  
Db 7796 GGTGCAAGGCCCTGGAGAACCGGATCAACGTGACCGTGTCTCAGTCCACACCGTGGTG 7855  
QY 8003 AATGCGAGGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGCACTGTGCTTGAAC 8062  
Db 7856 AACGCGAGGACTCGCAGGTTTCGCCGACGTGGAGTGCAGTTCGGTGCCTTGGCACTGCAT 7915  
QY 8063 ACAGCTACGGGCAACCGTTGGATGAGGAGAGGACAGGCTCCTGGAGCTGGCCCGCAG 8122  
Db 7916 GTGCGTATGGCATGACGCTGGACGAGGAGAAAGGCGCGCATTTCTGGAGCAGGCGCGCAG 7975  
QY 8123 AGAGCCGTGCGCCAAAGCGTGGGCCCGGAGCAGCAGAGACTGCGGGAAGGGAGGAAGGC 8182  
Db 7976 CGCGCGCTCGCCCGCGGTGGGCACGGGAGCAGCAGCGCTGCGCGACGCGGAGGAGGT 8035

QY	8183	CTGCGGGCCTGGACAGAGGGGAGAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAGGC	8244
Db	8036	CGCGCCCTCTGGACGGAGGGTGAGAAACGGCAGCTGCTGAGCGCTGGCAAGGTGCAGGC	8095
QY	8243	TACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTGACAGCGGCAAC	8302
Db	8096	TACGATGGGTACTACGTACTGTTCGGTGAGCAGTACCCGAGCTGGCTGACAGTGCCAAC	8155
QY	8303	AACATCCACTTTCATGAGACAGCGAGATGGGCCGAGGTGAC	8345
Db	8156	AACATCCAGTTCTTCGCACAAAGTGAGATCGGCACAGGTAAC	8198
RESULT 10			
AX662355	Sequence 37 from Patent WO02062999.		
LOCUS	AX662355	8645 bp	DNA linear PAT 22-MAR-2003
DEFINITION	AX662355		
ACCESSION	AX662355.1 GI:29163217		
VERSION	1		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X., Malyankar, U., Shinkets, R.A., Tchernev, V.T., Spaderna, S.K., Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A., Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F., Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G., Millet, I. and Macdougall, J.R.		
TITLE	Proteins and nucleic acids encoding same		
JOURNAL	Patent: WO 02062999-A 37 15-AUG-2002;		
FEATURES	Curagen Corporation (US)		
source	Location/Qualifiers		
	1..8645		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match 41.9%; Score 3502.2; DB 6; Length 8645;			
Best Local Similarity 65.3%; Pred. No. 0;			
Matches 5449; Conservative 0; Mismatches 2688; Indels 204; Gaps 13;			
QY	35	ATGGACGTGAAGGAGAGGAGCCTTACCGCTCGCTGACCCGGCGC---CGCGACGCCGAG	91
Db	151	ATGGATGTGAAGAACGCGAGCCCTTACTGCTCCCTGACCAAGACGACGAGAGAGAA	210
QY	92	CGCCGCTACACCGACTGTCCCGGACAGCGAGGAGGGCAAAG--CCCCGACAGAAATCG	148
Db	211	CGCGCTACACAAATTCCCTCGCAGACAATGAGGAGTGCCGGGTACCCACACAGAAATCC	270
QY	149	TACAGTCCAGCGAGACCTGTAAGGCCTACGACCAAGGA---CGCCCGCCTAGCCTATGCC	205
Db	271	TACAGTCCAGCGAGACATTGAAAGCTTTTGATCATGATTCCTCGGGCTGCTTTACGGC	330
QY	206	AGCCGCGTCAAGGACATTGTCCCGCAGGAGGCCGAGGAATCTGCCGACAGGTGCCAAC	265
Db	331	AACAGAGTGAAGGATTTGGTTTCAGAGAGCAGACGAGTTCACAGCAAGGACAGAAAT	390
QY	266	TTACCCCTGCGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGG	325
Db	391	TTTACCCCTAAGGCAGTTAGAGATTGTGAACCAAGCAACTCGAAGAGGACTGGCATTTGT	450
QY	326	ACAGACATTGGCCTGCCCAATGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATG	385
Db	451	GCGGAAATGGGGCTCCCTCACAGAGTTACTCTATCAGTCGAGGGTCAGATGCTGATACT	510
QY	386	GAGGCTGACACGGTGCTGTCCCTGAGCACCCCGTGGCTGTGTGGGGCCCGGACACAGG	445
Db	511	GAAAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTGGGCGAGGGGGTCAA	570



Db 3658 AGCGCAGCATTTCTCGCCCGCAGTTGCAATGGTCAAGCTGATGGTAACAAGTTACTGGCC 3717

QY 3728 CCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATT 3787

Db 3718 CCAGTGGCGCTAGCTTGTGGGATCGATGGCAGTCTGTACTAGCGGATTTTCAACTACGTG 3777

QY 3788 AGAAGATCTTCCCCTCTGGAATGTCAACCAATCTTAGAGCTGAGGAATAAAGATTTC 3847

Db 3778 CGCGGATATTCCCTTCTGGAATGTAAACAAGTGTCTTAGAACTAAGAAATAAAGATTTT 3837

QY 3848 AGACATAGTCAAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATAGTGGGGCC 3907

Db 3838 AGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCAACGGGAGAT 3897

QY 3908 GTCTTCCTTTCTGACAGCAACAGCCGGCGGGTCTTTAAATAAAGTCCACTGTGTGGTG 3967

Db 3898 CTGTACGTTTCTGACACAAACACCGCGCAGAAATTTATCGCCCAAGTCACTTTACGGGGCA 3957

QY 3968 AAGGACCTTGTCAAGAACTCTAGGTGGTTGCGGGGACAGGTGACCAAGTGCCTCCCTTT 4027

Db 3958 AAAGACTTGACTAAATAATGCAGAAAGTCTGCGAGGACAGGGGAGCAATGCCTTCCGTTT 4017

QY 4028 GATGACACTCGCTCGGGGATGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCCAGG 4087

Db 4018 GACGAGGCGAGATGTGGGATGGAGGAAGGCCGTGGAAGCCACACTCATGAGTCCCAA 4077

QY 4088 GGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCTGTTGATGGCACCATGATCAGACGC 4147

Db 4078 GGAATGSCAGTTGATAAGAAATGGATTAACTACTTCTTGTGATGAACCATGATTAGAAA 4137

QY 4148 ATCGATCAGAAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATCAGCCGG 4207

Db 4138 GTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTGACTTTCAGCCAGA 4197

QY 4208 CCACTCAGTGTGATTCTGTCTATGGATATTTCCCAGGTAAGACTGGAGTGGCCCCACAGAC 4267

Db 4198 CCTTTAACTGTGACACCAGCATGCACATCAGCCAGTACGTCTGGAATGGCCCCACTGAC 4257

QY 4268 TTAGCCATCAACCCAAATGGACAACCTCACTTTATGTCTCGACAACAATGTGGTCTCGCAA 4327

Db 4258 CTAGCCATTAACCCCTATGGATAACTCCATTTATGTCTCTGGATAATAATGATTTTACAG 4317

QY 4328 ATCTCTGAAACCCACAGGTGGCAATGTCTGCCGGAGGCCCATGCACCTGCCAGTCCCT 4387

Db 4318 ATCACTGAAATCGTCAAGTTTCGCAATTGCTGTGACGGCCCATGCACTGTCAAGGTTCCC 4377

QY 4388 GGCATTGACCACTTCTCTGCTAAGCAAGGTGGCCCATCCACGCAACCCCTGGAGTCAGCCACC 4447

Db 4378 GGAGTGG---AATATCTGTGGGAAGCAGCGGGTGCAGACAACACTGGAATCAGCCACT 4434

QY 4448 GCTTTGGCTGTTTACACAATGGGTCCTGTATATTGCTGAGACTGATGAGAAAAGATC 4507

Db 4435 GCCATTGCTGTCTCTACAGTGGGTCCTGTACATTACTGAACTGATGAGAAGAAAT 4494

QY 4508 AACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGCCCCCAGT 4567

Db 4495 AACCGGATAAGCAGGTCAACAACAGATGGAGAAATCTCCTTAGTGGCCGGAATACCTTCA 4554

QY 4568 GGCTGTGACTGTAAATAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTATGCC 4627

Db 4555 GAGTGTGACTGTCAAAAATGATGCCAACTGTGACTGTTTACCAGAGTGGAGATGGCTACGCC 4614

QY 4628 AAGGATGCAAGTTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGAGCTCTAC 4687

Db 4615 AAGGATGCCAAACTCAGTGGCCCCATCTCCTGGCTGCTTCTCCAGATGGTACACTGTAT 4674

QY 4688 GTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTCCTCAAC 4747

Db 4675 ATTGCAGATCTAGGGAATATCCGGATCCGGCTGTGTCAAGAATAAAGCCTTTACTTAAC 4734

QY 4748 ACCAGAACATGTATGAGCTGTCTTCAACAAATTGACGAGCTCTATCTGTTTGATACC 4807

Db 4735 TCTATGAACCTTCTATGAAGTTGCGTCTCCAACTGATCAAGAACTCTACATCTTTGACATC 4794

QY 4808 ACCGGCAAGCACCTGTATACACCCAAAGCCTGCCACAGGAGACTPACCTGTACAACCTTCACC 4867

Db 4795 AATGGTACTCAACCAATATACTACTGTAAAGTTTAGTCACTGGTGAATACCTTTACAATTTAGC 4854

QY 4868 TACACTGGGACGGCGACATCACTCATCACAGACAACAATGGCAACATGGTAAATGTC 4927

Db 4855 TACAGCAATGACAATGATATTACTGCTGTGACAGACAGCAATGGCAACACCCCTTAGAATT 4914

QY 4928 CGCCGAGACTCTACTGGGATGCCCTCTGGTGGTCCAGATGGCCAGGTGTACTGG 4987

Db 4915 AGACGGGACCCAAATCGCATGCGAGTTCGAGTGGTGTCTCCTGATAACCAAGTGATAIGG 4974

QY 4988 GTGACCATGGSCACCAACAGTGCACCTCAAGAGTGTGACCACACAAGGACACGAGTTGGCC 5047

Db 4975 TTGACAATAGGAACAAATGGATGTTTGAAAGGCATGACTGCTCAAGGACTGGAATTAGTT 5034

QY 5048 ATGATGACATACCATGGCAATTCGCGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGG 5107

Db 5035 TTGTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAGTGATGAAACTGGATGG 5094

QY 5108 ACAACATTTTATGAGTACGACAGCTTTGGCCGCTTGACAAAATGTGACCTTCCCTACTGGC 5167

Db 5095 ACAACGTTTTTGTACTATGACAGTGAAGTCTGCTGACAAAATGTACGTTTCCAACTGGA 5154

QY 5168 CAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGCTCCAGGTAGAGACCTCC 5227

Db 5155 GTGGTCAAAAACCTGCATGGGGACATGGACAAGGCTATCACAGTGGACATGAGTCATCT 5214

QY 5228 AGC-- --AAGGATGATGTCAACATAACCCAAACCTGTCTGCCCTCAGGGCCCTTCTACACA 5284

Db 5215 AGCCGAGAAGAAGATGTACGATCACTTCAAAATCTGTCTCGATCGATTCTTTCTACACC 5274

QY 5285 CTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCCTTGCCTGG 5344

Db 5275 ATGGTTCAAGATCAGTTAAGAAACAGCTACCAGATTGGTTATACGGCTCCCTCAGAAAT 5334

QY 5345 CTGCTGGCCAAACGGCATGGAGTGGCGTGCAGACTGAGCCCCACTGCTGGCTGGCACC 5404

Db 5335 ATCTACGCCAGTGGCCTGGACTCACACTACCAACAGAGCCGCACGTTCTGGCTGGCACC 5394

QY 5405 GTCAACCCACCGTGGGCAAGAGGAATGTACGCTGCCATCGACAACGGCCTCAACCTG 5464

Db 5395 GCTAATCCGACGTTGCCAAAAGAAACATGACTTTGCCCTGGCGAGAACGGTCAAAACTTG 5454

QY 5465 GTGAGTGGCGCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCGCCGG 5524

Db 5455 GTGGAATGGAGATTCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTCTTTGGCCGCAAG 5514

QY 5525 CTGCGGGTGCACAACCGAAATCTCCTATCTCTGGACTTTTGATCGCGTAACACGACAGAG 5584

Db 5515 CTCAGGGTTAATGGCAGAAACCTCCTTTTCACTGACTTTTGATCGAACAACAAGACAGAA 5574

QY 5585 AAGATCTATGATGACCAACCGCAAGTTTCACTTCCCTTCCGATTCTGTACGACCAAGCGGGCGG 5644

Db 5575 AAGATCTATGACGACCAACCGTAAATTTTCTACTGAGGATCGCCTACGACACGTCTGGGCAC 5634

QY 5645 CCCAGCCTCTGTGTCACCCAGCAGCGGCTGAATGGTGTCAACCGTGACATACTCCCTGGG 5704

Db 5635 CCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGGCCGTCAATGTCACTATTTCATCCACA 5694

QY 5705 GGTTACATTGCTGGCATCCAGAGGGGCTCATGTCTGAAAGAAATGGAATACGACACGGCG 5764

Db 5695 GGTCAAAATTGCCAGCATCCAGCGAGGCCACTAGCGAGAAAGTAGATTATGACGGACAG 5754

QY 5765 GGCCGCATCATCTCCAGGATCTTCTGCTGATGGTAAACATGGAGTTACACATATTTAGAA 5824

Db 5755 GGGAGGATCGTGTCTCGGGTCTTTGCTGATGGTAAACATGGAGTTACACATATTTAGAA 5814

QY 5825 AAGTCCCATGGTGTCTACTACACAGCCACAGGCGCAGTATATCTTTGAGTTCGACAAGAAT 5884

Db 5815 AAGTCCCATGGTTCTTCTGCTTCATAGCCAGCGGCAGTACATCTTCGAATACGATATGTGG 5874



Db	8017	GGCGCGTGGCGTGACAGTGGCGTACGGCATACCTGGACGAGGAGAGCGCGCATC	8076
QY	8105	CTGAGCTGGCCCGCAGAGAGCCGTGCGCCAAAGCGTGGGCCCGCAGCAGAGAGACTG	8164
Db	8077	CTGAGCAGGCGCGCAGCGCGCGCTCGCCCGGCGCTGGCGCGCAGCAGCAGCGCGTG	8136
QY	8165	CGGAAGGGGAGGAAGCCTCGGGGCCTGGACAGAGGGGAGAAAGCAGAGTGTGAGC	8224
Db	8137	CGCGACGCGAGGAGGCGCGCGCCTCTGGACGGAGGGCGAGAAAGCGCAGCTGTGAGC	8196
QY	8225	ACAGGCGGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCTGAGCAGTACCCAGAA	8284
Db	8197	GCCGCAAGGTGCAGGGCTACGACGGGTACTACGTACTCTCGGTGGAGCAGTACCCCGAG	8256
QY	8285	CTGTGAGACAGCGCCCAACAATCCACTTTCATGAGACAGAGCGAGATGGGCCGAGGTGA	8344
Db	8257	CTGGCGACAGCGCCCAACAATCCAGTTCTCTCGGCGAGAGCGAGATCGGCGAGGTAA	8316
QY	8345	C	8345
Db	8317	C	8317
RESULT 11			
AX662353			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 5445; Conservative			
QY	35	ATGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCCGCGC---CGCGACGCCGAG	91
Db	151	ATGGATGTGAAGAAGACGCAGGCCTTACTGCTCCCTGACCAAGAGCAGACGAGAGAGGAA	210
QY	92	CGCGCTACACCCAGCTCGTCCGCGGACAGCGAGGGGCAAG---CCCCGAGAAATCG	148
Db	211	CGCGCTACACAAATTCCTCCGACAGACAAATGAGGAGTGC CGGTACCCACACAGAGTCC	270
QY	149	TACAGCTCCAGCGAGACCTTGAAGGCCTACGACCAGGA---CGCCCGCTAGCCTATGCC	205
Db	271	TACAGTTCACCGAGACATTGAAGCTTTTGATCATGATTCTCTCGCGCTGCTTTACGCG	330
QY	206	AGCCCGCTCAGGACATTGTGCCGCGAGAGGCCGAGGAATTCGCGCACAGGTGCCAAC	265
Db	331	AACAGAGTGAAGATTGGTTTCACAGAGAAGCAGACGAGTTCACTAGACAAGGACAGAA	390
QY	266	TTACCCCTGCGGAGCTGGGCTGGAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGG	325

Db	391	TTTACCTAAGGCAGTTAGGAGTTTGTGAACCAGCAACTCGAAGAGGACTGGCATTTGT	450
QY	326	ACAGACATTGGCCTGCCCCAAATCGCGGCTACTCCATGGGGGCTGGCTCTGTATGCCGACATG	385
Db	451	GCGBAATGGGCTCCCTCACAGAGGTTACTCTATCAGTCAGGGTCAGATGCTGATACT	510
QY	386	GAGGTGACACGGTGTCTCCCTGAGCACCCCGTGGCTGTGTGGGCCGAGCAGCACGG	445
Db	511	GAATGAAGCAGTGATGTCCCGAGAGCATGCCATGAGACTTTGGGCGAGGGGTCAAA	570
QY	446	TCAGGCGCAGCTCCTGCTGTCCAGCCGGCCAAATTCAAATCTCACTACCCGACACC	505
Db	571	TCAGGCGCAGCTCCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACCTGACAGATACG	630
QY	506	GAGCATCAAAAACACTGAGACTGATCATCCGGCGGCTCGAGAAACACGCGGCTCCGG	565
Db	631	GAGCACGAAAAACAAGTCCGACAGTGAGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT	690
QY	566	ACGCGCGCGCGCGCTCTCGCAAGCCACACCCCCAAACAGACACACGCGGCTCCATT	625
Db	691	ACCCTGCAGCCCTTGC CGCTTCCCATAAAGCAGCACTCTGCACAGCATCATCCATCCATC	750
QY	626	AACCTCCCTGAACCGGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCCCCACGGAC	685
Db	751	ACTTCTCTCAACAGAAACTCCCTGACCAATAGAAGAACCCAGAGTCGGCCCCCGCGCT	810
QY	686	CACCTGCTCTCCGAGAGCCCCCTGCGCGCGGCCAGGAGCTGCCACGCCCAGGAG	745
Db	811	GCTTTGCCCGCGAGCTGCAAC-----CACACCCGAGTCGTCAGCTGCAGGAC	861
QY	746	AACCTGGCTGCTCAACAGCAACATCCCCCTGGAGACCAGGAACCTAGGCAAGCAGCCATTC	805
Db	862	AGCTGGGTCCTTGGCAGTAATGTACCACCTGGAAAGCAGGCCATT-----	905
QY	806	CTAGGGACATTGCAGGACAAACCTCATTGAGATGGACATTCTCGGCGCCTCCCGCCATGAT	865
Db	906	-----	905
QY	866	GGGCTTACAGTGACGGGCACTTCTCTCTTCAAGCCTGGAGGCACTCCCGCTCTTCTGC	925
Db	906	-----CCTATTCAAAAACAGGAACAGGTACAAAGCCACTGTTTCACT	945
QY	926	ACCACATCACAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCC	985
Db	946	ACTGCAACCCCGAGGATACACAATGGCATCTGGCTCTGTTTATTACACCACTACTCGGCCA	1005
QY	986	CTGCCCGCAGCACCTTCGCCCCGCGCGCTTTAAACCTCAAGAAGCCTCCAAGTACTGT	1045
Db	1006	CTACCTAGAAAACACCCCTATCAAGAAGTGCTTTAAATTCAGAAGTCTTCAAAGTACTGT	1065
QY	1046	AACCTGAAGTGGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCTCTGCTG	1105
Db	1066	AGCTGGAATGCACTGCACTGTGTGCGGTAGGGGTCTCGGTGCTCCTGGCAATACTCCTG	1125
QY	1106	GCATACTTTGTGGCCATGCACCTGTTTGGCTAAACTGGCACCTGCAGCCCGATGGAGGG	1165
Db	1126	TCTTATTTTATAGCAATGCATCTCTTTGGCTCAACTGGCAGCTACAGCAGACTGAAAAT	1185
QY	1166	CAGATGTATGAGATCACGGAGGACACAGCGAGCAGTTGGCCTGTGCCAACCGACGTCTCC	1225
Db	1186	GACACATTTGAGA-----ATGGAAGTGAATTCTGATACCATGCCAACAAACACTGTG	1239
QY	1226	CTATACCCCTCAGGGGCACCTGGCTTAGAGACCCCTGACAGGAAGGCAAGGAACCAACA	1285
Db	1240	TCATTACCTTTGAGACA-----	1258
QY	1286	GAAGGAAAGCCAGTAGTTTCTTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGAT	1345
Db	1259	-ATGGAATAATTAGTGGATTACGCAAGAAAAATAACACCATAGATTCCGGAGAACTTGAT	1317
QY	1346	GTGGGAAGCGAGCTCCAGAAAGATTCTCTCTGGCACTTTCTCTGGAGATCTCAAGTGTTC	1405
Db	1318	ATTGCCCGAAGAGCAATTCAAGAGATTCTCTCCGGGATCTTCTGGAGATCACAGCTCTTC	1377

QY	1406	ATAGACCATCCTGTGCATCTGAAATTC	AAATCAATGTGTCTCTGGGAAAGGCAGCCCTGGTTGGC	1466
Db	1378	ATTGATCAGCCACAGTTTCTTAAATTC	AATATCTCTCTCAGAAGGATGCATTGATTGGA	1437
QY	1466	ATTTATGGCAGAAAAGGCCTCCTCTT	TCACATA-----CACAGTTTGACTTT	1513
Db	1438	GTATATGGCCGGAAGATTACCGCCTT	TCCCATACTCAGTCCTCCCCCAGTATGACTTC	1497
QY	1514	GTGGAGCTGCTGGATGGCAGAGGCTC	CTAACCCAGGAGCGCGGAGCCTAGAGGGGACC	1573
Db	1498	GTGGAGCTCCTGGATGGCAGAGGCTG	ATTGCCAGAGACGAGCGCTGCTTGTGAGACG	1557
QY	1574	CCGCGCCAGTCTCGGGGA	CTGTGCCCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTAT	1633
Db	1558	GAGAGAGCCGGCGGCAGGCGAGATCC	GTGAGCCCTTTATGAGGCCGGCTTTATCCAGTAC	1617
QY	1634	TTGGATT	CAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGAGTCAGAACTGGT	1693
Db	1618	TTGGATTCTGGAATCTGGCACTGGCT	TTTTTAAATGATGGGAAAAATGCAGAGCAGGTG	1677
QY	1694	TCCTTTCTCACCACTGCCACTTGAGTC	CGGTGGATAACTGCCCCCAGCAACTGCTATGGCAAT	1753
Db	1678	TCCTTTAATACCAATTGTTATAGAGTC	TGTGGTGGAAATGTCCCCGAAATTGCCATGGAAAT	1737
QY	1754	GGTGACTGCATCTCTGGACCTGCCACT	GTCTTCTGGGTTTCTGGGCCCGACTGTGGC	1813
Db	1738	GGAGAAATGCGTTCTGGAACTTGCCAT	TGTTTCCAGGATTTCTGGTCCGGATTGTCA	1797
QY	1814	AGAGCCTCCTGCCCGCTGCTCTGTAG	CGGAAATGGCCAAATACATGAAAGCAGATGCTTG	1873
Db	1798	AGAGCCGCTGTCCAGTTGATGTAGTG	GCAACGGGAGTACTCCAAGGCGCGTGCCTG	1857
QY	1874	TGCCACAGTGGCTGGAAGGCGCTGAG	TGCGATGTCCTTCCACCAACAGTGTATCGATGTG	1933
Db	1858	TGTTTCAGCGGCTGGAAGGCGCACCG	AGTGTGATGTGTCGCGACTACCCAGTGTATGACCCA	1917
QY	1934	GCCTGCAGCAACCATGGCACTGCATC	ACGGGCACTGCACTGTCAACCCCTGGCTACAAG	1993
Db	1918	CAGTGTGGGGTCTGTGGATTGTATC	ATGGGCTCCTGTGCTTGCAGCTCAGATACAAA	1977
QY	1994	GGCGAGAGCTGTAGGAAGTGGACTGC	ATGGACCCACATGTTCAGGCGGSGTGTCTGC	2053
Db	1978	GGAGAAAGTTGTGAAGAAGCTGACTG	TATAGACCTGGGTGTTCTAATCATGGTGTGT	2037
QY	2054	GTGAGAGCGGAATGCOATTGCTTTGT	GGAATGGGAGGCACCAACTGCGACACCCCGAGG	2113
Db	2038	ATCCACGGGGAATGTCACTGCAGTCC	AGGATGGGAGGTAGCAATTGTGAATACTGAAG	2097
QY	2114	GCCACATGCTTAGACCAGTGTTCAGG	CCACCTTCTCCGGACACCGGCTTTGC	2173
Db	2098	ACCATGTGTCCAGACCAGTGCTCCGG	CCACCGAACGTATCTTCAAGAAAGTGGCTCCTGC	2157
QY	2174	AGCTGTGACCCCAAGCTGGACTGGACA	CGACTGTCTATCGAGATCTGTGCTGCCGACTGT	2233
Db	2158	ACGTGTGACCCCTAACTGGACTGGCC	CCAGACTGCTCAACGAAATATGTTCTGTGGACTGT	2217
QY	2234	GGTGGCCATGGCGTGTGCGTAGGGGG	CACTGCGCTGCGAGGATGGCTGGATGGGGCA	2293
Db	2218	GGCTCACACGGCGTTTGCATGGGGGG	ACGTGTCGTGTGAAGAAAGCTGGACGGGCCCA	2277
QY	2294	GCCTGCGACCAAGCGGCGCTGCCAC	CCCGCTGTGTGCCGAGCATGGGACCTGCGCGCAGCGC	2353
Db	2278	GCCTGTAATCAGAGAGCCTGCCACCC	CCCCGCTGTGTCGAGCACGGGACCTGCAAGGATGGC	2337
QY	2354	AAGTGCAGTGCAGCCCTGGCTGGAAT	GGGAACTGCACCACTGCACCATCGCTCACTATCTGGAT	2413
Db	2338	AAGTGTGAATGCAGCCAGGGCTGGA	ATGGAGAGCACTGCACATATCGCTCACTATTTGGAT	2397
QY	2414	AGGGTAGT-----TAAAGAGGGT	TGCCCTGGTGTGTGCAATGGCAAC	2455
Db	2398	AAGATAGTTAAAGACAAGATAGGATA	TAAAGAGGGTGTCTCTGTGCAACAGCAAT	2457

Qy	2456	GGCAGATGTACCTTTAGACCTGAATGGTTGGCACTGCGCTGTGCCAGCTGGCGTGGAGAGGA	2511
Db	2458	GGAAAGATGTACCCCTGGACCAAAAATGGCGGACATTTGTGTGTGCCAGCCTGGATGGAGAGGA	2517
Qy	2516	GCTGGCTGTGACACTTCCATGGAGACTGCCCTGCGGTGACAGCAAAAGACAATGATGGAGAT	2575
Db	2518	GCAGGCTGTGACGTAGCATGGAGACTCTTTGCACAGATAGCAAGGACAATGAAGGGGAT	2577
Qy	2576	GGCCTGGTGGACTGCATGGACCCCTGACTGCTGCCAGCCCTGTGCCATATCAACCCG	2635
Db	2578	GGACTCATTTGACTGCGATCCCGATTGCTGCTACAGAGTTCTCTGCCAGAATCAGCCC	2637
Qy	2636	CTGTGCCCTTGGCTCCCTAAACCCCTCTGGACATCATCCAGGAGACACAGSTCCCTGTGTCA	2695
Db	2638	TATTGTCGGGGACTGCCGGATCCTCAGGACATCATTAGCCAAAGCCTTCAATCGCCTTCT	2697
Qy	2696	CAGCAGAACCTACACTCCTTCTATGACCGCATCAAGTTCTCTGTTGGGAGGGACAGCACG	2755
Db	2698	CAGCAAGCTGCCAAATCCTTTTATGATCGAATCAGTTTCTTATAGGATCTGATAGCACC	2755
Qy	2756	CACATAATCCCCGGGGAGAAACCCCTTTTGATGGAGGGCATGCTTGTGTATTCTGTGGCCAA	2815
Db	2758	CATGTTATACCTGGAGAAAGTCCTTTTCAATAAGAGCCTTGCACTGTGTCATCAGAGGCCAA	2817
Qy	2816	GTGATGACATCAGATGGAAACCCCTTGGTTGGTGTGAACATCAGTTTTTGTCAATAACCTT	2875
Db	2818	GTACTGACTGCTGATGGAACTCCACTTATTGGAGTAAATGTCTCGTTTTTCCATTACCCA	2877
Qy	2876	CTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTTGACTTGGTGACAAATGGCGGC	2935
Db	2878	GAATATGGATATACTATTACCCGCCAGGACGGAATGTTTGACTTGGTGGCAAAATGGTGGG	2937
Qy	2936	ATCTCCATCATCCTGCGGTTTCGAGCGGGGACCTTTTCATCACACAGAGCACACCCTGTGG	2995
Db	2938	GCCTCTCTAACTTTGGTATTTGAAACGATCCCCATTCCTCACTCAGTATCATACTGTGTGG	2997
Qy	2996	CTGCCATGGGATCGCTTCTTTGTCTATGGAACCATCATCATGAGACATGAGGAGAAATGAG	3055
Db	2998	ATTCCATGGAAATGTCTTTTATGTGATGGATPACCTTAGTCCATGGAGAAAGAAGAAATGAC	3057
Qy	3056	ATTCCCAGCTGTGACCTGAGCAATTTTGCCCGCCCCCAACCCAGTGGTCTCTCCATCCCCA	3115
Db	3058	ATTCCCAGCTGTGATCTGAGTGAATTCGTGAGGGCCAAATCCCATCATTTGTGTATCACCT	3117
Qy	3116	CTGACGTCTTCGCCAGCTCCTGTGCGAGAGAAAGGCCCCCATTTGTGCGGAAATTCAGGCT	3175
Db	3118	TTATCCACCTTTTTCAGATCTTCTCCTGAAGACAGTCCCATCATTTCCGAAACACAGGTA	3177
Qy	3176	TTGCAGGAGGAAATCTCTATCTCTGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGG	3235
Db	3178	CTCCACGAGGAACTACAAATCCAGGAACAGATTTGAAACTCTCTACTTTGAGTTCCAGA	3237
Qy	3236	ACCCCTGGCTACAAATCTGTCTCTGAGGATCAGCCTCACCCACCCGACCATCCCCCTTCAAC	3295
Db	3238	GCTGCAGGGTATAAGTCAGTTCTCAAGATCACCATGACCCAGTCTATTATTCATTTAAT	3297
Qy	3296	CTCATGAAGGTGCACCTCATGGTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTTCGT	3355
Db	3298	TTAATGAAGGTTTCATCTTATGGTAGCTGTAGTGAAGACTCTTCCAAAAGTGGTTTCCT	3357
Qy	3356	GCAGCCCCAGACCTGTCCCTATTATTTCATTTGGGACAAGACAGACGCTCTACAACCAAG	3415
Db	3358	GCCTCACCAAACTTGGCCTATCTTCATATGGGATAAAACAGATGCATATAATCAGAAA	3417
Qy	3416	GTGTTTGGCTTTCAGAAAGCCTTTGTTTCCGTGGGTTATGAATATGAATCCTGCCCCAGAT	3475
Db	3418	GTCTATGGTCTATCTGAAGCTGTTGTCTCAGTTGGATATGAGTATGAGTCGTGTTTGGAC	3477
Qy	3476	CTAATCCTGTGGGAAAAAGAACACAGTGTCTGCAGGGCTATGAAATTCACGCGTCCAAG	3535
Db	3478	CTGACTCTGTGGGAAAAAGAGGACTGCCATTCTGCAGGGCTATGAATTTGGATGCGTCCAAC	3537
Qy	3536	CTTGGAGATGGAGCCTAGACAAACATCATGCCCTCAACATTCAAAAGTGGTATCCTGTCAC	3595

Db 3538 ATGGTGGCTGGACATTAGATAAAACATCACGTGCTGGATGTACAGAACGGTACTGTAC 3597  
QY 3596 AAAGGGAATGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCCTGTCTCATTTGGGAGCATCATG 3655  
Db 3598 AAGGGAACCGGGAAACCAAGTTTCAATCTCCAGCAGCCTCCAGTCGTGAGTAGCATCATG 3657  
QY 3656 GGCAATGGGCGCGGAGAGACATCTCTGCCCCAGCTGCAACGGCCTTGCTGACGGCAAC 3715  
Db 3658 GGCAATGGGCGAAGGCGCAGCATTTCTGCCCCAGTTGCAATGGTCAAGCTGATGGTAAC 3717  
QY 3716 AAGCTCCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGAT 3775  
Db 3718 AAGTTACTGGCCCCAGTGGCGCTAGCTTGTGGATCGATGGCAGTCTGTACGTAGGCGAT 3777  
QY 3776 TTCAACTACATTAGAAGGATCTTCCCTCTGGAAATGTCACCAACATCCTAGAGCTGAGG 3835  
Db 3778 TTCAACTACGTGGCGGATATTCCCTTCTGGAATGTAACAAGTGTCTTAGAATAAGA 3837  
QY 3836 AATAAAGATTTCAGACATAGTCACAGTCCAGCACACAATACTACTACCTGGCCACAGACCCC 3895  
Db 3838 AATAAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACTCTTGCAACGATCCA 3897  
QY 3896 ATGAGTGGGCGCTCTTCCCTTTCTGACAGCAACAGCCCGGGTCTTTTAAATCAAGTCC 3955  
Db 3898 GTCACGGGAGATCTGTACGTTTCTGCACAAACACCCGAGAAATTTATCGCCCAAGTCA 3957  
QY 3956 ACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTCGGGGACAGGTGACCCAG 4015  
Db 3958 CTTACGGGGCAAAAGACTTGTACTAAAAATGCAGAAATCGTCGCAGGGACAGGGAGCAA 4017  
QY 4016 TGCCTCCCTTTTGATGACACTCGCTCGGGGGATGGTGGGAAGGCCACAGAAAGCCACTC 4075  
Db 4018 TGCCTTCCGTTTGACGAGGCGAGATGTGGGGATGGAGGGAAGGCCCGTGGAAAGCCACTC 4077  
QY 4076 ACCAATCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGTGGATGGCACC 4135  
Db 4078 ATGAGTCCCAAGGAATGGCAGTTGATAAAGATGGATTATCTACTTTTGTGATGGAACC 4137  
QY 4136 ATGATCAGACGCATCGATCAGATGGGATCATCTCCACCCTGCTGGCTCTAATGATCTC 4195  
Db 4138 ATGATTAGGAAAGTTGACCAAAATGGAATCATATCAACTCTCTTGGGCTCTAACGATTTG 4197  
QY 4196 ACATCAGCGCGCCACTCAGCTGTGATTCTGTCTATGGATAATTTCCAGGTAAGACTGGAG 4255  
Db 4198 ACTTCAGCCAGACCTTTAACTTGTGACACCAGCATGCACATCAGCCAGGTACGTCCTGGAA 4257  
QY 4256 TGGCCCAAGACTTAGCCATCAACCAATGGACAACCTCACTTTATGTCTCGACAAACAAT 4315  
Db 4258 TGGCCCACTGACCTAGCCATTAACTTATGGATAACTCCATTTATGTCTCGATAATAAT 4317  
QY 4316 GTGGTCTCGAAATCTCTGAAACACCAAGGTGCGCATTTGTGCGCGGGAGGCCCATGAC 4375  
Db 4318 GTAGTTTACAGATCACTGAAATCGTCAAGTTGCAATTTGCTGCTGGACGGCCCATGAC 4377  
QY 4376 TGGCAGGTCCTGGCATTGACCACCTTCTGCTAAGCAAGGTGGCCATCCACGCCAACCCCTG 4435  
Db 4378 TGTCAAGTTCGCGAGTGG--AATATCTGTGGGGAAGCACCGGGTGCAGACAACACTG 4434  
QY 4436 GAGTCAGCCACCGCTTTGGCTGTTTCACACAATGGGGTCCGTATATTTGCTGAGACTGAT 4495  
Db 4435 GAATCAGCCACTGCCATTGCTGTCTACAGTGGGGTCTGTACATTACTGAAACTGAT 4494  
QY 4496 GAGAAAAAGATCAACCGCATCAGGCAAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCT 4555  
Db 4495 GAGAGAAAAATTAAACCGGATAAGGCAAGGTCAACAACAGATGGAGAAATCTCCTTAGTGGCC 4554  
QY 4556 GGGGCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTTCTGGAGAC 4615  
Db 4555 GGAATACCTTCAGAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTTACCAGAGTGA 4614  
QY 4616 GATGGTTATGCCAAGGATGCAAGTTAAATACCCCATCTTCCCTGGCTGTGTGTGCTGAT 4675

Db 4615 GATGGTACGCCCAAGGATGCCAAACTCAGTGCCCCATCTCTCCCTGGCTGCTTCTCCAGAT 4674  
QY 4676 GGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAG 4735  
Db 4675 GGTACACTGTATATTGCAGATCTAGGGAATATCCGGATCCGGCTGTGTCAAAGAATAAG 4734  
QY 4736 CCTTTCTCAACACACCCAGAACATGTATGAGCTGTCTTTCACCAATTGACCAGGAGCTCTAT 4795  
Db 4735 CCTTTACTTAACCTCTATGAACCTTCTATGAAGTTGCGTCTCCAATGATCAAGAATCTAC 4794  
QY 4796 CTGTTTGATACCAACCGGCAAGCCTGTACACCCAAAGCCTGCCCCACAGGAGACTACCTG 4855  
Db 4795 ATCTTTGACATCAATGGTACTCTCACAATATACTGTAAAGTTTAGTCACTGGTGATTACCTT 4854  
QY 4856 TACAACCTTCACTACTGGGGACGGCGACATCACACTCATCACAGACAACAATGGCAAC 4915  
Db 4855 TACAATTTTAGCTACAGCAATGACAATGATATTACTGCTGTGACAGACAGCAATGGCAAC 4914  
QY 4916 ATGGTAAATGTCGCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGC 4975  
Db 4915 ACCCTTAGAATTAGACGGGACCCAAATCGCATGCCAGTTTCGAGTGGTGTCTCTCTGATAAC 4974  
QY 4976 CAGGTGTACTGGGTGACCATGGGCAACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGA 5035  
Db 4975 CAAGTGATATGGTTGACAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGA 5034  
QY 5036 CACGAGTTGGCCATGATGACATACCATGGCAATTCGGGCCTTCTGGCAACCAAAAGCAAT 5095  
Db 5035 CTGGAATTAGTTTGTTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAAGTGAT 5094  
QY 5096 GAAAAACGGATGGACAACAATTTTATGAGTACGACAGCTTTTGGCCGCTGACAAAATGTGACC 5155  
Db 5095 GAAACTGGATGGACAACGTTTTTTGACTATGACAGTGAAGGTGCTGTGACAAAATGTTACG 5154  
QY 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAAGTGATACAGACAGTTCAAGTGCATGTCCAG 5215  
Db 5155 TTTCCAACCTGGAGTGGTCACAAACCTGTCATGGGACATGGACAAGGCTATCACAGTGGAC 5214  
QY 5216 GTAGAGACCTCCAGC--AAGGATGATGTCAACCATAAACCACCAACCTGTCTGCCCTCAGGC 5272  
Db 5215 ATTGAGTCACTAGCCGAGAAGAAGTGTGAGCATCACTTCAAATCTGTCTCTCGATCGAT 5274  
QY 5273 GCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCCGATGGC 5332  
Db 5275 TCTTTCTACACCATGTTTCAAGATCAGTTAAGAAACAGCTACCAGATTGTTATGACGGC 5334  
QY 5333 TCCTTGGCGCTGCTGTGGCCCAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTG 5392  
Db 5335 TCCCTCAGAAATTATCTACGCCAGTGGCTGGACTCACACTACCAACAGAGCCGACGTT 5394  
QY 5393 CTGGTGGCACCGTCAACCCCAACCGTGGCAAGAGGAATGTCAAGCTGCCCATCGACAAC 5452  
Db 5395 CTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAACATGACTTTGCCTGGCGAGAAC 5454  
QY 5453 GGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGAGAGGCTCGGGGCCAGGTCACTGTC 5512  
Db 5455 GGTCAAAACCTTGGTGGAAATGGAGATTCCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTC 5514  
QY 5513 TTTGGGCGCGGCTGGGGTGCACAACCCGAAATCTCCTATCTCTGGACTTTGATCGCGTA 5572  
Db 5515 TTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTTTCAGTTGACTTTGATCGAACA 5574  
QY 5573 ACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTCAACCTTCGGATTTCTGACGAC 5632  
Db 5575 ACAAAGACAGAAAAAGATCTATGACGACCAACCGTAAATTTCTACTGAGGATCGCCTACGAC 5634  
QY 5633 CAGGCGGGCGGCCAGCCTCTGTCCACCGAGCAGCAGGCTGAATGTTTCAACGTGACA 5692  
Db 5635 ACGTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGGCGGTCAATGTCAACC 5694  
QY 5693 TACTCCCCCTGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAAGAAATGGAA 5752  
Db 5695 TATTTCATCCACAGTCAAATTTGCCAGCATCCAGCGAGGCACCACTAGCGAGAAAGTAGAT 5754

QY	5753	TACGACCAAGCGGGCCGCATCACATCCAGGATCTTTCGCTGATGGGAAGACATGGAGCTAC	5811
Db	5755	TATGACCGACAGGGGAGGATCGTGTCTCGGGTCTTTGCTGATGGTAAACATGGAGTTAC	5814
QY	5813	ACATACTTAGAGAAGTCCATGGTGTCTACTACACAGCCAGAGGCAGTATATCTTTGAG	5872
Db	5815	ACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGCCAGCGGCAGTACATCTTCGAA	5874
QY	5873	TTTCGACAAGAAATGACCGCCTCTCTTCTGTGACGATGCCCAACGTGGCGCGGCACACACTA	5932
Db	5875	TACGATATGTGGACCGCCTGTCTGCCATCACCATGCCCAGTGTGGCTCGCCACACCATG	5934
QY	5933	GAGACCATCCGCTCAGTGGGCTACTACAGAAAACATCTATCAGCCCCCTGAGGGGCAATGCC	5992
Db	5935	CAGACCATCCGATCCATTGGCTACTACCGCAACATATACAAACCCCGGAAAGCAACGCC	5994
QY	5993	TCAGTCATACAGGACTTCACTGAGGATGGCAACCTCTTCAACACCTTCTACCTGGGCACCT	6052
Db	5995	TCCATCATCACGGACTACACGAGGAAGGGCTGCTTCTACAAAACAGCTTCTTGGGTACA	6054
QY	6053	GGCCGCAAGGTGATATACAAAGTATGGCAAACTGTCAAGCTGGCAGAGACGCTCTATGAC	6112
Db	6055	AGTCGGAGGGTCTTATTCAAATACAGAAAGGCAGACTAGGCTCTCAGAAAATTTTATATGAT	6114
QY	6113	ACCACCAAGGTCAGTTTCACTATGACGAGACGGCAGGCATGCTGAAGACCAACCTA	6172
Db	6115	AGCACAAAGATCAGTTTACCTATGATGAACAGCAGGAGTCTTAAAGACAGTAAACCTC	6174
QY	6173	CAGAAATGAGGGCTTCACTGCACCATCCGCTACCGTCAGATTTGGGCCCTGATTGACCGA	6232
Db	6175	CAGAGTGATGGTTTTATTGACCATTAGATACAGGCAAAATGGTCCCTGATTGACAGG	6234
QY	6233	CAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTTGGACTACAACCTATGAC	6292
Db	6235	CAGATTTTCCGCTTTAGTGAAGATGGGATGGTAAATGCAAGATTTGACTATAGCTATGAC	6294
QY	6293	AACAGCTTCCGGGTGACCAAGCATGCAGGCTGTGATCAACGAGACCCCACTGCCCAATTGAT	6352
Db	6295	AACAGCTTTCGAGTGACCAAGCATGCAGGGTGTGATCAATGAACGCCCACTGCCCTATTGAT	6354
QY	6353	CTCTATCGCTATGATGTGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTATT	6412
Db	6355	CTGTATCAGTTTGATGACATTTCTGGCAAAAGTTGAGCAGTTTGGAAAAGTTTGGAGTTATA	6414
QY	6413	TACTATGACATTAACCAAGATCATCACCAAGCTGTGATCAACGAGACCCCAACCAAGCATTTTGTAT	6472
Db	6415	TATTATGATATTAAACCAAGATCATTTCTACAGCTGTGAATGACCTATACGAAGCAGCTTTGTAT	6474
QY	6473	GCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATG	6532
Db	6475	GCTCATGGCCGTATCAAGAGATTTCAATATGAGATATTCAAGTCTGCTCATGTACTGGATT	6534
QY	6533	ACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTAC	6592
Db	6535	ACAATTCAGTATGATAACATGGGTCCGGTAAACCAAGACAGAGATTAAATATAGGGCCCTTT	6594
QY	6593	GCCAAATACCACTCGCTACTCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCTCC	6652
Db	6595	GCCAAACACCACCAATATGCTTATGAATATGATGTTGATGGACAGCTCCAAACAGTTTTAC	6654
QY	6653	ATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACCTTACTG	6712
Db	6655	CTCAATGAAAAGATAATGTGGCGGTACAACCTACCATCTGAATGGAAACCTCCATTTACTG	6714
QY	6713	AGCCCTGGGAACAGTGCAGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACT	6772
Db	6715	AACCCAAAGTAACAGTCGCGCTGTGACACCCCTTCGCTATGACCTCGGAGACAGAATCACT	6774
QY	6773	CGGCTGGGTGACGTGCAATACAAGATGGATGGATGGCTTCTTGAGGACAGCGGGCGGT	6832
Db	6775	CGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGTTTTCTACGTCAAGAGGGGCACG	6834

QY	6833	GATATCTTTTGAGTACAACTCAGCTGGCCCTGCTCATCAAGGCCTACAAACCGGCTGGCAGC	6899
Db			
QY	6835	GAAATCTTTTGAATATAGCTCCAAGGGCTTCTAACTCGAGTTTACAGTAAAGCGAGTGGC	6894
Db			
QY	6893	TGGAGTGTCCAGGTACCGTACGATGGCCCTGGGGCGCGCGTGTCCAGCAAGACAGCCAC	6952
Db			
QY	6895	TGGACAGTGATCTACCGTTATACGGCTTGGGAAGCGGTGTTTCTAGCAAAACCAAGTCTA	6954
Db			
QY	6953	AGCCACCACCTGCAGTTCTTCTATGCAGACCTGACCAACCCCAACCAAGGTACCCACCTG	7012
Db			
QY	6955	GGACAGCACCTGCAGTTTTTTATGCTGACTTAACTTATCCCACTAGGATTAAGTCAATGTC	7014
Db			
QY	7013	TACAAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTT	7072
Db			
QY	7015	TACAAACCATTCGAGTTCAGAAATTACCTCCCTGTATTATGATCTCCAAGGACATCTTTT	7074
Db			
QY	7073	GCCATGGAGCTGAGCAGTGGTGATGAGTTTATACATAGCTTTGTGACAAACATCGGGACCCCT	7132
Db			
QY	7075	GCCATGGAAATCAGCAGTGGGATGAAATTCATATATTGCATCGGATAACACAGGGACACCA	7134
Db			
QY	7133	CTTGCTGCTTTAGTGGAAACAGGTTTGATGATCAAGCAAATCCTGTACACAGCCTATGGG	7192
Db			
QY	7135	CTGGCTGTGTTTCAGTAGCAATGGGCTTATGCTGAAACAGATTCAGTACACTGCATATGGG	7194
Db			
QY	7193	GAGATCTACATGGATACCAACCCCAACCTTTCAGATCATCATAGGCTACCATGGTGGCCTC	7252
Db			
QY	7195	GAAATCTATTTTGACTCTAAATATTGACTTTTCAACTGGTAATTGGATTTCAATGGTGGCCTG	7254
Db			
QY	7253	TATGATCCACTCACCAAGCTTGTCCACATGGGCGCGGAGATTATGATGCTGCTGGCCGGA	7312
Db			
QY	7255	TATGACCCACTCACCAAAATTAAATCCACTTTGGAGAAAAGAGATTATGACATTTTGGCAGGA	7314
Db			
QY	7313	CGCTGGACTAGCCGAGACCAACGAGCTGTGGAAGCACTTTAGTAGCAGCAACGTATGCCT	7372
Db			
QY	7315	CGGTGGACAACACCTGACATAGAAATCTGGAA--AAAGAAATTGGGAAGGACCCAGTCCCT	7371
Db			
QY	7373	TTTAATCTCTATATGTTTCAAAAACAACAACCCCATCAGCAACTCCCAAGACATCAAGTGC	7432
Db			
QY	7372	TTTAACTTGTTACATGTTTAGGAATAACAACCCCTGCAAGCAAAATCCATGACGTGAAAGAT	7431
Db			
QY	7433	TTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTGATC	7492
Db			
QY	7432	TACATCACAGATGTTAAACAGCTGGCTGGTGACATTTGGTTTCCATCTGCACAATGCTATT	7491
Db			
QY	7493	CCTGTTTATCCAAAACCAAGACATGGATGCCATGGAACCCCTCCTACGAGCTCATCCACACA	7552
Db			
QY	7492	CCTGGATTCCCTGTTCCCAAATTTGATTTAAACAGAACTTCTTACGA-----A	7539
Db			
QY	7553	CAGATGAAAACCGCAGGAGTGGGACAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTA	7612
Db			
QY	7540	CTTGTGAAGAGTCAGCAGTGGGATGATATACCGCCCATCTTTCGGAGTCCAGCAGCAAGTG	7599
Db			
QY	7613	CAGAAGCAGCTCAAGGCCCTTTGTCACTTAGAACGGTTTGCACCAGCTCTATGGCTCCACA	7672
Db			
QY	7600	CGCGGCAGGCCAAGGCCCTTCCCTGTCGCTGGGGAAGATGGCCGAGGTGCAGGTGAGCCGG	7659
Db			
QY	7673	ATCACCAGTCTCCAGCAGGCTCCAAAAGACCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTT	7732
Db			
QY	7660	CGCCGGGCCGGC--GGCGCGCAGTCCCTGGCTGTGGTTTCGCCACGGTCAAGTGCCTGATC	7716
Db			
QY	7733	GGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTG	7792
Db			
QY	7717	GGCAAGGGGTTCATGCTGGCCGTTCAGCCAGGGCCCGGTGCAGACCAACGTGCTCAACATC	7776
Db			
QY	7793	GCCAAATGAGGATGGGCGAAGGTTTGCTGCCATCTTGAACCATGCTGCCCATACCTAGAGAAC	7852
Db			
QY	7777	GCCAAACGAGGACTGCATCAAGGTGGCGGCCGTGCTCAACAAACGCTTCTACCTGGAGAAC	7836
Db			
QY	7853	CTGCATTACCAATTGATGGGGTGGATACCCATTACTTTGTGAAACCAAGGACCTTTCAGAA	7912
Db			
QY	7837	CTGCATTTACCAATCGAGGGCAAGGACACGCACACTACTTCATCAAGACCAACCGCCCGAG	7896
Db			
QY	7913	GGTGACCTTGCCCATCTCTGGGCCCTCAGTGGGGGGCGGCGAACCCTCGGAGAAATGGGTCAAC	7972
Db			

Db 7897 AGCGACCTGGGCACGCTGCGGTTGACAGCGGCCGCAAGGCGTGGAGAACGGCATCAAC 7956  
QY 7973 GTCACCTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATC 8032  
Db 7957 GTGACGGTGTGCGCAGTCCACACGCTGGTGAACGGCAGGACGGCAGGTTGCGCGGCGTG 8016  
QY 8033 CAGCTCCAGTACGGGGCAGTGTGCTTGAACACACAGCTACGGGACAAACGTTGGATGAGGAG 8092  
Db 8017 GAGATGCACTTCGGCGCGCTGGCGGTGACAGTGCGCTACGCGATGACCTGGACAGGAG 8076  
QY 8093 AAGGCACGGGTCTCTGGAGCTGGCCGGCAGAGAGCGGTGCGCCAAAGCGTGGGCCCGGAG 8152  
Db 8077 AAGGCGCGCATCTTGAGCAGGCGCGGAGCGCGGCTGCGCCGGGCTGGCGCGCGGAG 8136  
QY 8153 CAGCAGAGACTGCGGGAAGGGAGGAAGGCGCTGCGGGCCCTGGACAGAGGGGGAGAGCAG 8212  
Db 8137 CAGCAGCGGTGCGCAGCGCGGAGGAGGCGCGCCTCTGGACGAGGGCGAGAGCGG 8196  
QY 8213 CAGGTGTGAGCACAGGGCGGGTGAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAG 8272  
Db 8197 CAGCTGTGAGCGCGCGCAAGGTGACGGGCTACGACGGGTACTACGTACTCTCGGTGGAG 8256  
QY 8273 CAGTACCCAGAACTGTCAGACAGCGCCCAACACATCCACTTCATGACAGACAGAGCGAGATG 8332  
Db 8257 CAGTACCCGAGCTGGCCGACAGCGCCCAACACATCCAGTTCCTGCGGCAGAGCGAGATC 8316  
QY 8333 GGCCGAGAGTGAC 8345  
Db 8317 GGCAGGAGGTAAC 8329

RESULT 12  
AX662357  
LOCUS AX662357 8473 bp DNA linear PAT 22-MAR-2003  
DEFINITION Sequence 39 from Patent WO02062999.  
ACCESSION AX662357  
VERSION AX662357.1 GI:29163218  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,  
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,  
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,  
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,  
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,  
Millet,I. and Macdougall,J.R.  
TITLE Proteins and nucleic acids encoding same  
JOURNAL Patent: WO 02062999-A 39 15-AUG-2002;  
Curagen Corporation (US)  
FEATURES  
source location/Qualifiers  
1..8473  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 41.0%; Score 3422; DB 6; Length 8473;  
Best Local Similarity 67.1%; Pred. No. 0;  
Matches 5027; Conservative 0; Mismatches 2370; Indels 93; Gaps 9;  
QY 880 CGGGCATTCTCTTCAAGCCTGG--AGGCACCTCCCGCTCTTCTGCACCATCACC 936  
Db 725 CAGGCATTCTCTATTCAAACAGGAACAGGTACAACGCCACTGTTCACTACTGCAACCCC 784  
QY 937 AGGGTACCCACTGACGTCCAGCACAGTGTACTCTCCTCCGCCCGACCCCTGCCCGCAG 996  
Db 785 AGGATACACAATGGCATCTGGCTCTGTTTATTCAACCACTACTCGGCCACTACCTAGAAA 844  
QY 997 CACCTTCGCCCGCGCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAGTG 1056

Db 845 CACCCCTATCAAGAAGTGCTTTTAAATTCAAGAAGTCTTCAAAGTACTGTAGCTGGAATG 904  
QY 1057 CGCAGCCCTGAGCGCCCATCGTCATCTCAGCCCACTCTGGTCACTCCTGCTGGCATACTTTGT 1116  
Db 905 CACTGCACTGTGCGCGTAGGGGTCTCGGTGCTCCTGGCAATACTCCTGTCTTATTTAT 964  
QY 1117 GGCCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTATGA 1176  
Db 965 AGCAATGCATCTCTTTGGCCCTCAACTGGCAGCTACAGCAGACTGAAAATGACACATTTGA 1024  
QY 1177 GATCACGGAGGACACAGCCAGCAGTTGGCCCTGTGCCAACCCGACGTCTCCCTATACCCCTC 1236  
Db 1025 GA-----ATGAAAAGTGAATTCTGATACCATGCCCCAACAACTGTGTCTATTACCTTC 1078  
QY 1237 AGGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGBACCAACAGAGGAAAGCC 1296  
Db 1079 TGGAGACA-----ATGGAAAATT 1096  
QY 1297 CAGTAGTTTCTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGGAAGCG 1356  
Db 1097 AGGTGGATTTACGCAAGAAAATAACACCATAGATTCCGGAGAACTTGATATTGGCCGAAG 1156  
QY 1357 AGCCTCCCAGAAAGATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416  
Db 1157 AGCAATTCAGAGATTCTCTCCCGGATCTTCTGGAGATCACAGCTCTTCAATTGATCAGCC 1216  
QY 1417 TGTGCATCTGAAATTCATGTGTCTCTGGGAAAGGACGCCCTGGTGGCAATTTATGGCAG 1476  
Db 1217 ACAGTTTCTTAAATTCATATCTCTCTCAGAAAGGATGCATTGATGGAGTATATGGCCG 1276  
QY 1477 AAAAGGCTCCCTCCTTACATACACAGTTTGAATTTGTGGAGTGTGGATGGCAGGAG 1536  
Db 1277 GAAAGGCTTACCGCCTTCCCATACTCAGTATGACTTCGTGGAGCTCCTGGATGGCAGCAG 1336  
QY 1537 GCTCCTAACCCAGGAGCGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGT 1596  
Db 1337 GCTGATTGCCAGAGAGCAGCGGAGCCTGCTTGAGACGAGAGCGGGCGGCGGAG 1396  
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTCAAGGAATCTGGCACTT 1656  
Db 1397 ATCCGTGAGCCTTCATGAGCGCGCTTATCCAGTACTTGGATTCTGGAATCTGGCATCT 1456  
QY 1657 GGCTTTTACAAATGACGGAAGGAGTGCAGAAAGTGGTTTCTCTTCTCACCCTGCCATTGA 1716  
Db 1457 GGCTTTTATAATGATGGGAAAATGACAGACAGGTGTCTTTTAAATACCATTTGTTATAGA 1516  
QY 1717 GTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGAATGCTGATCTCTGGGACCTG 1776  
Db 1517 GTCGTGGTGGAAATGTCCCCGAAATGCCATGGAATGGAGAAATCGGTTTCTGGAACTTG 1576  
QY 1777 CCACTGCTTCTCGGGTTTCTCGGGCCCGGACTGTGGCAGAGCCCTCCTGCCCGTGTCTG 1836  
Db 1577 CCATTGTTTCCAGGATTTCTGGTCCGATTTGTTCAAGAGCCGCTGTCCAGTGTATTG 1636  
QY 1837 TAGCGGAATGGCCAAATACATGAAGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGC 1896  
Db 1637 TAGTGGCAACGGGCAGTACTCCAAGGCGCTGCCTGTGTTCAGCGGCTGGAAGGCGAC 1696  
QY 1897 TGAGTGGATGTGCCCAACCAACCAAGTGTATCGATGTGGCTGACAGCAACCATGGCACCTG 1956  
Db 1697 CGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGGTCTGTGGGATTG 1756  
QY 1957 CATCACGGGCACCTGCATCTGCAACCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGA 2016  
Db 1757 TATCATGGGCTCTGTGTGTGCAACTCAGGATACAAAGGAAAAAGTTGTGAAGAAGCTGA 1816  
QY 2017 CTGCATGGACCCACATGTTTCAGGCGGGGTGTCTGCTGAGAGGCGAATGCCATTGCTT 2076  
Db 1817 CTGTATAGACCTGGGTGTCTTAATCATGTGTGTGTATCCACGGGGAATGTCACTGCAG 1876  
QY 2077 TGTGGGATGGGAGGCACCAACTGCGAGACCCCGGCGCCACATGCTTAGACCACTGTTT 2136

Db 1877 TCCAGGATGGGAGGTAGCAATTGTGAAATACTGAAGACCATTGTCTCCAGACCAGTGTCTC 1936  
QY 2137 AGGCCACGGAAACCTTCCCTCCCGGACACCCGGCTTTTGACGCTGTGACCCCAAGCTGGACTGG 2196  
Db 1937 CGGCCACGGAAACGTATCTTCAAGAAAGTGGCTCCCTGCACGCTGTGACCCCTAACTGGACTGG 1996  
QY 2197 ACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCCATGGCGTGTGCGTAGG 2256  
Db 1997 CCAGACTGCTCAACACGAAATATGTTCTGTGGACTGTGGCTCACACGGCGCTTTCATGGG 2056  
QY 2257 GGGACCTGCCGCTCGGAGGATGGCTGGATGGGGGACGCTGCCAGCCAGCGGGCCTGCCA 2316  
Db 2057 GGGACGCTGCTGCTGTGAAGAGGCTGGACGGGCCAGCCTGTAATCAGAGACCTGCCA 2116  
QY 2317 CCCGCGCTGTCCGAGCATGGGACCTGCCCGGACGCGCAAGTGGAGTGGAGCCCTGGCTG 2376  
Db 2117 CCCCCGCTGTCCGAGCACGGGACCTGCAAGGATGGCAAGTGTGAATGACGCCAGGGCTG 2176  
QY 2377 GAATGGCGAACACTGCCACCATCGCTCACTATCTGGATAGGGTAGT----- 2421  
Db 2177 GAATGGAGACACTGCACTATCGCTCACTATTTGGATAAGATAGTTAAAGACAAGATAGG 2236  
QY 2422 ---TAAAGAGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACTTAGACCTGAA 2478  
Db 2237 ATATAAGAGGGTTGCTCTGCTGTGCAACAGCAATGGAAGATGTACCTGGACCAAAA 2296  
QY 2479 TGGTGGCACTGGCTGTCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGA 2538  
Db 2297 TGGCGGACATTGTGTGCCAGCCTGGATGGAGGAGCAGGCTGTGACGTAGCCATGGA 2356  
QY 2539 GACTGCCCTGCGGTGACAGCAAGACAAATGATGGAGATGGCCCTGGTGGACTGCATGGACCC 2598  
Db 2357 GACTCTTTGCACAGATAGCAAGGACAAATGAAGGGGATGGACTCATTGACTGCATGGATCC 2416  
QY 2599 TGACTGTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCCCTTGGCTCCCTAACCC 2658  
Db 2417 CGATTGCTGCCTACAGATTCTGCCAGAAATCAGCCCTATTGTCCGGGACTGCCGGATCC 2476  
QY 2659 TCTGGACATCATCCAGGAGACACAGGTCCCTGTGTCAACAGCAGAACCTACACTCCTTCTA 2718  
Db 2477 TCAGGACATCATTAGCCAAAAGCCTTCAATCGCCTTCTCAGCAAGCTGCCAAAATCCTTTA 2536  
QY 2719 TGACCGCATCAAGTTCCTCGTGGGACAGGACAGCAGCACATAATCCCGGGGAGAACCC 2778  
Db 2537 TGATCGAATCAGTTTCCCTTATAGGATCTGATAGCACCCCATGTTATACCTGGAGAAATCC 2596  
QY 2779 CTTTGATGGAGGCGCATGTTGTGTTATTCGTGSCAAAGTGTGACATCAGATGGAACCCC 2838  
Db 2597 TTTCAATAAGAGCCTTGCACTGTGTCATCAGAGGCCAAGTACTGACTGTGATGGAATCC 2656  
QY 2839 CCTGGTTGGTGAACATCAGTTTGTGCAATAACCCCTCTCTTTGGATATACAATCAGCAG 2898  
Db 2657 ACTTATTGGAGTAATGCTCGTTTTTCCATTACCCAGAAATATGGATATACTATTACCCG 2716  
QY 2899 GCAAGATGGCAGCTTTGACTTGGTGACAAAATGGCGGCATCTCCATCATCTCGGGTTCTGA 2958  
Db 2717 CCAGGACGGAATGTTTGACTTGGTGCAAAATGGTGGGCGCTCTCTAACTTTGGTATTGA 2776  
QY 2959 GCGGGCACCTTTTCATCACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTTCTTTGT 3018  
Db 2777 ACGATCCCCATTCCTCACTCAGTATCATACTGTGTGGATTCATGGAAATGCTTTTATGT 2836  
QY 3019 CATGGAACCATCATATGAGACATGAGGAGAAATGAGATTCCCAGCTGTGACCTGAGCAA 3078  
Db 2837 GATGGATACCCCTAGTCATGGAGAAAGAGAAATGACATTCACGCTGTGATCTGAGTGG 2896  
QY 3079 TTTTGGCCGCCAACCCAGTCTCTCTCCATCCCCACTGACGCTCCTTCGCCAGCTCCTG 3138  
Db 2897 ATTCTGTAGGCCAAAATCCCATCATTTGTGTCACTACCTTTATCCACCTTTTTCAGATCTTC 2956  
QY 3139 TGCAGAGAAAGGCCCCCATTTGTGCCGAAATTCAGGCTTTTGAGGAGGAAATCTCTATCTC 3198  
Db 2957 TCCTGAAGACAGTCCCATCATTCCTCCGAAACACAGGTACTCCACGAGGAAACTACAATTC 3016

QY 3199 TGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCCTGGCTACAAAATCTGTCTCT 3258  
Db 3017 AGGAACAGATTTGAAACTCTCCTACTTGAGTTCCAGAGCTGCAGGGTATAAGTCAGTTCT 3076  
QY 3259 GAGGATCAGCCTCACCCACCCGACCATCCCCTTCAACCTCATGAAGTGCACCTCATGGT 3318  
Db 3077 CAAGATCACCATGACCCAGTCTATTATTCCATTTAATTAATGAAGTTCATCTTATGGT 3136  
QY 3319 ACCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTTATTA 3378  
Db 3137 AGCTGTAGTAGGAAGACTCTTCCAAAAGTGGTTCTCCTCAACCAAACTTGGCCTATAC 3196  
QY 3379 TTTTCATTTGGGACAAAGACAGACGCTCTACAAACAGAAAGTGTTTTGGGCTTTTCAGAAAGCCTT 3438  
Db 3197 TTTTCATATGGGATAAAACAGATGCATATATAATCAGAAAAGTCTATGGTCTATCTGAAGCTGT 3256  
QY 3439 TGTTCCTGGTGTATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAAC 3498  
Db 3257 TGTGTCAAGTTGGATATGAGTATGAGTCGTGTTGGACCTGACTCTGTGGGAAAAAGAGGAC 3316  
QY 3499 AACAGTGTGCAGGGCTATGAAATTTGACGCGTCCAAAGCTTGGAGGATGGAGCCTAGACAA 3558  
Db 3317 TGCCATTCTGCAGGGCTATGAATTTGGATGCGTCCAAACATGGTGGCTGACATTAGATAA 3376  
QY 3559 ACATCATGCCCTCAACATTCAAAGTGGTATCCTGTCACAAAGGAAATGGGAGAAACCAAGTT 3618  
Db 3377 ACATCAGTGTGATGTACAGAACGGTATACTGTACAGGGAACACGGGAAAAACCAAGTT 3436  
QY 3619 TGTGTCTCAGCAGCCTCTGTCTATTGGGAGGATCATGGCAATGGGCGCGGAGAAAGCAT 3678  
Db 3437 CATCTCCAGCAGCCTCCAGTCGTGAGTAGATCATGGCAATGGCGAAGCGGAGCGCAGCAT 3496  
QY 3679 CTCCTGCCCCAGCTGCAACGGCCTTGTGTCAGCGCAACAAGCTCCTGSCCCAGTGGCCCT 3738  
Db 3497 TTCCTGCCCCAGTTGCAATGGTCAAGCTGATGTTAACAAGTTACTGSCCCAGTGGCGCT 3556  
QY 3739 CACTGTGGCTCTGACGGGAGCCTCTATGTGGTGTATTTCAACTACATTAGAAGGATCTT 3798  
Db 3557 AGCTTGTGGGATCGATGGCAGTCTGTACGTAGGCGATTTCAACTACGTGCGGCGGATATT 3616  
QY 3799 CCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCA 3858  
Db 3617 CCCTTCTGGAATGTAAACAAGTGTCTTAGAACTAAGAAATAAAGATTTTAGACATAGCAG 3676  
QY 3859 CAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGGGCGCGTCTTCTCTTC 3918  
Db 3677 CAACCCAGCTCATAGATACTACCTTGCACACGATCCAGTCACGGAGATCTGTACGTTTC 3736  
QY 3919 TGACAGCAACAGCCGCGGTCTTTAAATACTCAAGTCCACTGTGTGTGTGAAGGACCTTGT 3978  
Db 3737 TGACACAAAACACCCGACAGAAATTTATCGCCCAAAGTCACTTACGGGGGCAAAAGACTTGAC 3796  
QY 3979 CAAGAACTCTGAGTGGTTGCGGGACAGGTCAGGTCAGTGCCTCCCTTTTGTGATGACACTCG 4038  
Db 3797 TAAAAATGCAGAAAGTCTCGCAGGGACAGGGGAGCAATGCCCTTCCGTTTGACGAGGCGAG 3856  
QY 4039 CTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCAGGGGTATTACAGT 4098  
Db 3857 ATGTGGGATGGAGGAAGGCCGTGGAAGCCACACTCATGAGTCCCAAGGAATGGCAGT 3916  
QY 4099 GGACAAGTTTGGGCTGATCTACTTCTGTGGATGGCACCATGATCAGACCCATCGATCAGAA 4158  
Db 3917 TGATAAGAAATGGATTAAATCTACTTTTGTGTGATGGAACCATGATTAGGAAGATTGACCAAAA 3976  
QY 4159 TGGGATCATCTCCACCTGTCTCGGCTCTAATGATCTCACATCAGCCCGCCACTCAGCTG 4218  
Db 3977 TGAATCATATCAACTCTTCTGGCTCTAACGATTTGACTTCAGCCAGACCTTTTAACTTG 4036  
QY 4219 TGATTCTGTGATGATATTTCCAGGTAAGACTGGAGTGGCCCCACAGACTTAGCCATCAA 4278  
Db 4037 TGACACCAGCATGCACATCAGCCAGGTACGTCTGTGAATGGCCCACTGACCTAGCCATTAA 4096

QY 4279 CCAATGGACAACCTCACTTTATGTCCTCGACAACAATGTGGTCTGCAAAATCTCTGAAAA 4338  
Db 4097 CCCTATGGATAAATCCATTTATGTCCTGGATAAATAATGTAGTTTACAGATCACTGAAAA 4156  
QY 4339 CCACCAAGTGGCATTGTGCGCGGAGGCCCATGCACTGCCAGGTCCCTGGCATTGACCA 4398  
Db 4157 TCGTCAAGTTCGCATTGCTGCTGGACGGCCCATGCACTGTGAGTTTCCGGAGTGG--A 4213  
QY 4399 CTTCCTGCTAAGCAAGGTGGCCCATCCACGCAACCCCTGGAGTGCAGCCACCGCTTTGSGTGT 4458  
Db 4214 ATATCCTGTGGGAAGCAGCGGTGAGACAACACTGGAATCAGCCACTGCCATTGCTGT 4273  
QY 4459 TTACACAAATGGGGTCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCCGCATCAG 4518  
Db 4274 GTCCTACAGTGGGTCTGTACATTACTGAAACTGATGAGAAAAATTAACCCGATAAG 4333  
QY 4519 GCAGGTACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCCCCCAGTGGCTGTGACTG 4578  
Db 4334 GCAGGTACAACAGATGGAGAAATCTCCTTAGTGGCCGGAATACCTTCAGAGTGTGACTG 4393  
QY 4579 TAAAAATGATGCCAACTGTGATTGTTTTTCTGGAGACGATGTTATGCCAAGGATGCAAA 4638  
Db 4394 CAAAAATGATGCCAACTGTGACTGTTACCAGAGTGGAGATGGCTACGCCAAGGATGCCAA 4453  
QY 4639 GTTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGAGCTCTACGTGGCCGACCT 4698  
Db 4454 ACTCAGTGCCTCCATCTCCTCGCTGCTTCTCCAGATGGTACACTGTATATTGCAGATCT 4513  
QY 4699 TGGGAACATCCGAATTCGGTTTTATCCGGAAGAACAAAGCCCTTCTCAACACCCGAACAT 4758  
Db 4514 AGGGAATATCCGGATCCGGCTGTGTCAAAGAATAAGCCTTACTTAACTCTATGAACCT 4573  
QY 4759 GTATGAGTGTCTTCACCAATTGACCAAGGAGCTCTATCTGTTTGATACCACCCGCAAGCA 4818  
Db 4574 CTAAGAATGTGCGTCTCCAACTGATCAAGAACTCTACATCTTTGACATCAATGTGACTCA 4633  
QY 4819 CCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACCTACACTGGGA 4878  
Db 4634 CCAATATACTGTAAGTTTAGTCACTGGTGATTACCTTTACAAATTTAGCTACAGCAATGA 4693  
QY 4879 CGCGACATCACACTCATCACAGACAACAATGGCAACATGGTAAATGTCCGCCGAGACTC 4938  
Db 4694 CAATGATATTACTGCTGTGACAGACAGCAATGGCAACACCTTAGAATTAGACGGGACCC 4753  
QY 4939 TACTGGGATGCCCTCTGCTGGTGGTGGTCCAGATGGCCAGGTGACTGTTGGTGACCATGGG 4998  
Db 4754 AATCGATGCCAGTTCGAGTGGTGTCTCCTGTATAACCAAGTGATGGTTGACAAATAGG 4813  
QY 4999 CACCAACAGTGCACCTCAAGAGTGTGACCACACAAGGACAGAGTTGGCCATGATGACATA 5058  
Db 4814 AACAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAATTAGTTTGTACTTA 4873  
QY 5059 CCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAACCGATGGACAAATTTTA 5118  
Db 4874 CCATGGCAATAGTGGCCTTTTAGCCACTAAAAGTGATGAAACTGGATGGACAAACGTTTTT 4933  
QY 5119 TGAGTACGACAGCTTTGGCCGCTGTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAG 5178  
Db 4934 TGACTATGACAGTGAAGTGTGCTGACAAATGTTACGTTTCCAACTGGAGTGGTCAAAA 4993  
QY 5179 TTTCCGAAGTGATACAGACAGTTCAAGTGCATGTCCAGTGAAGACCTCCAGC---AAGGA 5235  
Db 4994 CCTGCATGGGACATGGACAAGGCTATCACAGTGGACATGAGTCACTAGCCGAGAAGA 5053  
QY 5236 TGATGTACCATAAACCACCACTGTCTGCCTCAGGCGCTTCTACACACTGCTGCAAGA 5295  
Db 5054 AGATGTACAGCATCACTTCAAATCTGTCCTCGATCGATTCTTTCTACACCATGGTTCAGA 5113  
QY 5296 CCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCCTTGGCTGTCTGCTGGCCAA 5355  
Db 5114 TCAGTTAAGAAACAGCTACCAAGATTGGTTATGACGGCTCCCTCAGAATTATCTACGCCAG 5173  
QY 5356 CGGCATGAGGTGGCGCTGCAGACTGAGCCCCCACTTGTGGCTGGCACCCGTCAACCCAC 5415

Db 5174 TGGCCTGGACTCACACTACCAACAGAGCCGACGTTCTGGCTGGCACCGCTAATCCGAC 5233  
QY 5416 CGTGGGCAAGAGGAATGTCACTGCTGCCATCGAACAGGCTCAACCTGGTGGAGTGGCG 5475  
Db 5234 GGTGTCCAAAAGAAACATGACTTTGCTTGGCGAGAACGGTCAAAACTTGGTGAATGGAG 5293  
QY 5476 CCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGTCTTTGGGCGCGGCTCGGGTGCA 5535  
Db 5294 ATTCCGAAAAGAGCAAGCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAA 5353  
QY 5536 CAACCGAAATCTCCTATCTCTGGACTTTGATCGCTAACACGCACAGAGAAGATCTATGA 5595  
Db 5354 TGGCAGAAACCTCCTTTAGTTGACTTTGATCGAACCAACAAAGACAGAAAAAGATCTATGA 5413  
QY 5596 TGACCACCGCAAGTTCACCCCTTCGGATTCTGTACGACGAGGCGGGCGGCCAGCCTCTG 5655  
Db 5414 CGACCACCGTAAATTTCTACTGAGGATCGCCTACGACACGCTCTGGSCACCCGACTCTCTG 5473  
QY 5656 GTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGTTACATTGC 5715  
Db 5474 GCTGCCAAGCAGCAAGCTGATGGCCGTCAATGTCACTTATTCACAGGTCAAAATGC 5533  
QY 5716 TGGCATCCAGAGGGGCATCATGTCTGAAAGAATGGAATACGACAGGCGGGCCGCATCAC 5775  
Db 5534 CAGCATCCAGCGAGGCCACCACTAGCGAGAAAAGTAGATTATGACGACAGGGAGGATCGT 5593  
QY 5776 ATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATCTTAGAGAAGTCCATGGT 5835  
Db 5594 GTCTCGGGTCTTTGCTGATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGT 5653  
QY 5836 GCTGCTACTACAGCCAGAGGCGAGTATATCTTTGAGTTTCGACAAGAATGACCGCTCTC 5895  
Db 5654 TCTTCTGCTTCATAGCCAGCGGCACTACATCTTCGAATACGATATGTGGGACCGCTGTC 5713  
QY 5896 TTCTGTGACGATGCCCCAAACGTGGCGGCGCAGACACTAGAGACCACTCCGCTCAGTGGGCTA 5955  
Db 5714 TGCCATCACCATGCCAGTGTGGTCCGCCACACCATGCAGACCATCCGATCCATTGGCTA 5773  
QY 5956 CTACAGAAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAATACAGGACTTCACTGA 6015  
Db 5774 CTACCGCAACATATACAACCCCGGAAAGCAACGCCCTCCATCATCACGGACTACAACGA 5833  
QY 6016 GGATGGCACTCCTTTCACACCTTTCTACCTGGGCACTGGCCGAGGGTGATATACAAGTA 6075  
Db 5834 GGAAGGGCTGCTTCTACAAACAGCTTTCTTGGGTACAAGTCGAGGGTCTTATTCAAATA 5893  
QY 6076 TGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCCACCAAGGTTCAGTTTCACTTA 6135  
Db 5894 CAGAAGGCAGACTAGGCTCTCAGAAATTTTATATGATAGCACAAGAGTCAGTTTACCTTA 5953  
QY 6136 TGACGAGACGGCAGGCATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACCTGCAC 6195  
Db 5954 TGATGAACAGCAGGAGTCTCTAAAGACAGTAACCTCCAGAGTGATGGTTTTATTGTCAC 6013  
QY 6196 CATCCGCTACCGTCAGATTGGGGCCCTTGATTGACCGACAGATCTTCCGCTTCACTGAGGA 6255  
Db 6014 CATTAGATACAGGCAAAATTGGTCCCTGATTGACAGGCGAGATTTCCGCTTTAGTGAAGA 6073  
QY 6256 AGGCATGGTCAACGCCCGTTTTTGTACTAACACTATGACAACAGCTTCCGGGTGACCAAGCAT 6315  
Db 6074 TGGGATGGTAAATGCAAGATTTGACTATAGCTATGACAACAGCTTTCGAGTGACCAAGCAT 6133  
QY 6316 GCAGGTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGTGTCT 6375  
Db 6134 GCAGGTGTGATCAATGAAACGCCCACTGCCCTATTGATCTGTATCAGTTTGTATGACATTTT 6193  
QY 6376 AGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCACTTTACTATGACATTAACCAAGATCAT 6435  
Db 6194 TGGCAAAAGTTGAGCAGTTTGGAAAAGTTTGGAGTTATATATATGATATTAACCAAGATCAT 6253  
QY 6436 CACCACAGTGTCTATGACCCACACCAAGCACTTTTGTATGCTATGTCAGGATGAAGGAAGT 6495

Db 6254 TTCTACAGCTGTAATGACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGAT 6313  
Qy 6496 GCAGTATGAGATCTTCGGCTCGCTCATGTACTGGATGACCGTCCAGATGATAAATGAG 6555  
Db 6314 TCAATATGAGATATTCAGGTGCTCATGTACTGGATTACAATTCAAGTATGATAAATGAG 6373  
Qy 6556 GCGAGTAGTGAAGAGAGAGCTGAAGGTAGGAGCCCTACGCCAATACCACTCGCTACTCCTA 6615  
Db 6374 TCGGGTAACCAAGAGAGAGATTAAATAGGGCCCTTTGCCAACACCAATATGCTTA 6433  
Qy 6616 TGAGTATGATGCTGACGCGCCAGCTGCAGACAGTCTCCATCAATGACAGCCACTCTGGCG 6675  
Db 6434 TGAATATGATGTTGATGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCG 6493  
Qy 6676 CTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCACGGCT 6735  
Db 6494 GTACAACTACGATCTGAATGGAACCTCCATTACTGAACCCAAAGTAACAGTGCAGCTCT 6553  
Qy 6736 CACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGGTGCATACAA 6795  
Db 6554 GACACCCCTTCGCTATGACCTGCGAGACAGAACTCACTGACTGGGTGATGTTCAATATCG 6613  
Qy 6796 GATGGATGAGGATGGCTTCCTGAGGCGCGGGCGGTGATATCTTTGAGTACAACTCAGC 6855  
Db 6614 GTTGGATGAAGATGTTTCCTACGTCAAAGGGGACGGAATCTTTGAATATAGCTCCAA 6673  
Qy 6856 TGGCCTGCTCATCAAGGCTTACAAACCGGCTGGAGCTGGAGTGTCAAGTACCGCTACGA 6915  
Db 6674 GGGCTTCTAACTCGAGTTTACAGTAAAGGCAAGTGGCTGGACAGTGTCTACCGTTATGA 6733  
Qy 6916 TGGCCTGGGCGCGCGTGTCCAGCAAGAGCAGCAGCACAGCCACCACTGCGAGTTCTTCTA 6975  
Db 6734 CGGCTGGGAAGCGGTGTTCTAGCAAAACCAAGTCTAGGACAGCACCTGCGAGTTTCTTA 6793  
Qy 6976 TGCAGACCTGACCAACCCACCAAGGTCAACCCACTGTACAAACCACTCAGCTCTGAGAT 7035  
Db 6794 TGCTGACTTAATATCCCACTAGGATTACTCATGTCTACAACCACTTCAGATTCAAAAT 6853  
Qy 7036 CACCTCCCTCTACTACGACTTGAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGA 7095  
Db 6854 TACCTCCCTGTATTATGATCTCCAAGGACATCTTTGCCATGGAATCAGCAGTGGGA 6913  
Qy 7096 TGAGTTTACATAGCTTGTGACAAACATCGGGACCCCTCTTGTCTCTTTAGTGAACAGG 7155  
Db 6914 TGAATCTATATTGCAATCGGATAACACAGGGACACCACTGGCTGTGTTAGTAGCAATGG 6973  
Qy 7156 TTTGATGATCAAGCAATCCTGTACACAGCCCTATGGGAGATCTACATGATACCAACCC 7215  
Db 6974 GCTTATGCTGAAACAGATTCAGTACACTGCAATATGGGAAATCTATTTGACTCTAATAT 7033  
Qy 7216 CAACTTTACATCATATAGGCTACCATGTTGGCTCTATGATCCACTACCAAGCTTGT 7275  
Db 7034 TGACTTTCAACTGGTAATGGATTTTCAATGGTGGCTGTATGACCCACTCACCAATTAAT 7093  
Qy 7276 CCACATGGGCGCGGAGATTATGATGTCTGGCGGACGCTGGACTAGCCAGACACCA 7335  
Db 7094 CCACTTTGGAGAAAGAGATTATGACATTTTGGCAGGACGCTGGACAAACCTGACATAGA 7153  
Qy 7336 GCTGTGAAGCACCTTAGTAGCAGCAACGTCATGCTTTTAATCTCTATATGTTTCAAAA 7395  
Db 7154 AATCTGGA--AAAGAAATGGGAAGGACCCAGCTCCTTTTAACTTGTATGTTTAGGAA 7210  
Qy 7396 CAACAAACCCATCAGCAACTCCAGGACATCAAGTGTCTCATGACAGATGTTAAACAGCTG 7455  
Db 7211 TAACAAACCCCTGCAAGCAAAATCCATGACGTGAAAGATTACATCACAGATGTTAAACAGCTG 7270  
Qy 7456 GCTGCTCACCTTTGGATTCCAGCTACACAACGTCATGCTCCCTGGTTATCCCAAAACAGACAT 7515  
Db 7271 GCTGGTGACATTTGGTTTCCATCTGCACAATGCTATCTCTGGATTCCCTGTTTCCAAAT 7330  
Qy 7516 GGATGCCATGGAACCCCTCTACGAGCTCATCCACACACAGATGAAACACGAGGAGTGGGA 7575  
Db 7331 TGATTTAACAGAACCTTCTTACGA-----ACTTGTGAAGAGTTCAGCAGTGGGA 7378

Qy 7576 CAACAGCAAGTCTATCTCTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGT 7635  
Db 7379 TGATATACCGCCCATCTTCGGAGTCCAGACCAAGTGGCGCGGCAAGGCCCTTCCT 7438  
Qy 7636 CACCTTAGAACCGTTTGACAGCTCTATGGCTCCACAATCACCAAGTCCAGCAGGCTCC 7695  
Db 7439 GTCGCTGGGAAGATGGCCGAGGTGCAGGTGAGCCGGCCGGCCGGC--GGCGCGCA 7495  
Qy 7696 AAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGGTCAAGTTTGCCTT 7755  
Db 7496 GTCCTGGCTGTGTTTCGCCACGGTCAAGTCTGATCGGCAAGGGGCTCATGCTGGCCGT 7555  
Qy 7756 GAAGGATGGCCGAGTGACACACAGACATCATCAGTGTGGCCAATGAGATGGCGGAAGGT 7815  
Db 7556 CAGCCAGGGCCGCTGCAGACCAACGTCGTCAACATCGCCAACGAGGACTGCATCAAGGT 7615  
Qy 7816 TGCTGCCATCTTGAACCATGCCCACTACCTAGAGAACTTGACATTCACCATTTGATGGGT 7875  
Db 7616 GCGGGCCGTGCTCAACAACGCTTCTACCTGGAGAACCTGCACATTCACCATCGAGGCAA 7675  
Qy 7876 GGATACCCATTTACTTTGTGAACACAGGACCTTCAGAAAGGTGACCTGGCCATCCTGGCCCT 7935  
Db 7676 GGACAGCACTACTTCATCAAGACCAACCGGCTGAGAGCGACCTGGGACGCTGCGGTT 7735  
Qy 7936 CAGTGGGGGGCGGAACCTTGAGAAATGGGGTCAACGTCACGTGTCGCCAGATCAACAC 7995  
Db 7736 GACCAGCGCGCAAGCGCTGGAGAACCGCATCAACGTGACGTTGTCGAGTCCACCAC 7795  
Qy 7996 AGTACTTAATGGCAGACTAGACGCTACACACACATCCAGCTCCAGTCCGGGCACTGTG 8055  
Db 7796 GGTGGTGAACGGCAGGACCGCAGGTTTCGGGACGTGGAGATGACGTCGGCGCTGGC 7855  
Qy 8056 CTTGAACACACGCTACGGGACAAACGTTGGATGAGGAGAAAGGCAAGGTCCTGGAGCTGC 8115  
Db 7856 GCTGACAGTGGCTACGGCATGACCCCTGGACGAGGAGAAAGCGCGCATCTTGAGCAGGC 7915  
Qy 8116 CCGCAGAGAGCGCTGCGCCCAAGCGTGGGCGCGGAGCAGCAGAGACTGCGGGAAGGGA 8175  
Db 7916 GCGCAGCGCGCGCTGCGCGCGGCTGGGCGCGGAGCAGCAGCGCTGCGCGCAGCGA 7975  
Qy 8176 GGAAGGCTGCGGGCTGGACAGAGGGGGAAGCAGAGCTGCTGACACAGGCGCGGT 8235  
Db 7976 GGAGGGCGCGGCTCTGACGAGGGCGGAGAAAGCGGAGCTGCTGAGCGCGGCAAGGT 8035  
Qy 8236 GCAAGGCTACGACGCTTTTCTGATCTCTGTCAGCAGTACCCAGAACTGTCAGACAG 8295  
Db 8036 GCAGGGCTACGACGCGGTACTACGTACTCTCGTGGAGCAGTACCCCGAGCTGGCCGACAG 8095  
Qy 8296 CGCCAAACAACTCCACTTCATGAGACAGAGCGGAGATGGCCGAGGTGAC 8345  
Db 8096 CGCCAAACAACTCCAGTTCCTGCGGAGAGCGGAGATCGGACGAGGTAAAC 8145

RESULT 13  
AX662359  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

AX662359  
Sequence 41 from Patent WO02062999.  
AX662359  
AX662359.1 GI:29163219

8487 bp  
DNA  
linear  
PAT 22-MAR-2003

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,  
Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,  
Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangoli, E.A.,  
Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,  
Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,  
Millet, I. and Macdougall, J.R.  
Proteins and nucleic acids encoding same



Db 2671 CTCGTTTTTCCATTACCAGAAATATGGATATACTATTACCCGCCAGGACGGAATGTTTGA 2730  
QY 2917 CTTGGTGACAAATGGCGGCATCTCCATCATCTCGCGTTCGAGCGGCACCTTTTCATCAC 2976  
Db 2731 CTTGGTGGCAAATGGTGGGCTCTCTAACTTTGGTATTGTAACGATCCCCCATTCCTCAC 2790  
QY 2977 ACAGGAGCACACCCTGTGGCTGCCATGGGATCGCTTCTTTGTTCATGGAAACCATCATCAT 3036  
Db 2791 TCAGTATCATACTGTGTGGATTCCATGGAAATGTCCTTTATGTGATGGATACCCTAGTCAT 2850  
QY 3037 GAGACATGAGGAGAAATGAGATTCCCAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCC 3096  
Db 2851 GGAGAAAGAGAAATGACATTCCCAGCTGTGATCTGAGTGGATTCTGTGAGGCCAAATCC 2910  
QY 3097 AGTCGTCTCTCCATCCCACCTGACCTCCTTCGSCAGCTCCTGTGCGAGAAAGGCCCCAT 3156  
Db 2911 CATCATTTGTCTCATCACCTTTATCCACCTTTTCAGATCTTCTCCTGAAGACAGTCCCAT 2970  
QY 3157 TGTGCGGAAATTCAGGCTTTGCAGGAGGAATCTCTATCTCTGCTGCGAAGATGAGGCT 3216  
Db 2971 CATTTCCGAAACACAGGTACTCCACGAGGAATACTAATTTCCAGGAACAGATTTGAAACT 3030  
QY 3217 GAGCTACCTGAGCAGCCGACCCCTGGCTACAAATCTGTCTCCTGAGGATCAGCCTCACCCA 3276  
Db 3031 CTCCTACTTGAGTTCAGAGCTGCAGGGTATAAGTCAGTTCTCAAGATCACCATGACCCA 3090  
QY 3277 CCGGACCATCCCTTCAACCTCATGAAGGTGCACTCATGCTGAGTGGTGGAGGGCCGCT 3336  
Db 3091 GTCTATTATTCATTTAATTAATGAAGGTTTCATCTTATGTTAGTCTGTAGTGAAGACT 3150  
QY 3337 CTTCAGGAAGTGGTTCGCTGCAGCCGCCAGACCTGTCTCTATTATTTCAATTTGGGACAAGAC 3396  
Db 3151 CTTCCAAAAGTGGTTTCTCGCTCACCAAACTTGGCCTATACTTTTCATATGGGATAAAAC 3210  
QY 3397 AGAGCTCTACAACAGAGGTGTTTGGGCTTTTCAGAAAGCCTTTGTTTCCGTGGTTATGA 3456  
Db 3211 AGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTCACTGGATATGA 3270  
QY 3457 ATATGAATCCTGCCCAGATCTAATCCTGTGGGAAAGAAACAAACAGTGTGAGGGCTA 3516  
Db 3271 GTATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAGAGGACTGCCATTTCTGAGGGCTA 3330  
QY 3517 TGAATTTGACGCTCCAAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCTCAACAT 3576  
Db 3331 TGAATTTGATCGTCCAAACATGGTGGCTGGACATTAGATAAACATCACGTGCTGATGT 3390  
QY 3577 TCAAAGTGGTATCCTGACAAAGGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCC 3636  
Db 3391 ACAGAACGGTATACTGTACAGGGAACCGGGAAACCAAGTTCTATCTCCACAGCCTCC 3450  
QY 3637 TGTCAATGGGAGCATCATGGCAATGGCGCCGGAGAGCATCTCCTGCCCCAGCTGCAA 3696  
Db 3451 AGTCGTAGTAGCATCATGGGCAATGGCGAAGGCGGAGCATTTCTGCCCCAGTTGCAA 3510  
QY 3697 CGGCTTGTGACGGCAACAGCTCCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGG 3756  
Db 3511 TGGTCAAGCTGATGGTAACAAGTTACTGGCCCCAGTGGCGCTAGCTTTGTGGGATCGATGG 3570  
QY 3757 GAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTCCCCTCTGGAAATGTCAC 3816  
Db 3571 CAGTCTGTACGTAGGCGATTTCAACTACGTGGCGGGATATTTCCCTTTCTGGAAATGTAAC 3630  
QY 3817 CAACATCCTAGAGCTGAGGAATAAGATTTTCAGACATAGTCACAGTCCAGCACACAAATA 3876  
Db 3631 AAGTGTCTTAGAACTAAGAAATAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATA 3690  
QY 3877 CTACCTGGCCACAGACCCCATGAGTGGGCGCTCTTCTCTTTCTGACAGCAACAGCCGCG 3936  
Db 3691 CTACCTTGCAACGGATCCAGTCCAGGAGATCTGTACGTTTCTGACACAAACACCCGCGAG 3750  
QY 3937 GGTCCTTAAATCAAGTCCACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGT 3996

Db 3751 AATTTATCGCCCCAAAGTCACTTACGGGGSCAAAAGACTTGTACTAAAAATGCAGAAGTCGT 3810  
QY 3997 TCGGGGACAGGTGACCAAGTGCCTCCCTTTTGATGACACTCGCTGGGGGATGGTGGAA 4056  
Db 3811 CGCAGGGACAGGGAGCAATGCCTTCCGTTTGACGAGGCGAGATGTGGGGATGGAGGAA 3870  
QY 4057 GGCCACAGAAGCCCACTCAACCAATCCAGGGGTATTACAGTGGCAAGTTTGGGCTGAT 4116  
Db 3871 GGCCGTGGAAGCCCACTCATGATGCCAAAGGAATGGCAGTTGATAAAGAAATGGATTAAT 3930  
QY 4117 CTACTTCGTGGATGGCACCATGATCAGACGCATCGATCAGAAATGGGATCATCTCCACCCT 4176  
Db 3931 CTACTTTGTTGATGGAACCATGATTAGGAAAGTTGACCAAAATGGAATCATATCAACTCT 3990  
QY 4177 GCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGTGATTTCTGTCTATGGATAT 4236  
Db 3991 TCTGGGCTCTAACGATTTGACTTCAGCCAGACCTTTAACTTTGTGACACAGCATGCACAT 4050  
QY 4237 TTCCAGGTAAGACTGGAGTGGCCCAACAGACTTAGCCATCAACCCCAATGGACAACTCACT 4296  
Db 4051 CAGCCAGGTACGTCCTGGAATGGCCCACTGACCTAGCCATTAAACCCCTATGGATAACTCCAT 4110  
QY 4297 TTATGTCCTCGACAAACAAATGTGCTCCTGCAAAATCTCTGAAAACCAACAGGTGCGCATTTGT 4356  
Db 4111 TTATGTCCTGGATAATAATGATGTTTACAGATCACTGAAAATCGTCAAGTTTCGATTTGC 4170  
QY 4357 CGCCGGAGGCCCATGCACTGCCAGTCCCTGGCATTGACCACTTCTCTGCTTAAGCAAGGT 4416  
Db 4171 TGCTGGACGGCCCATGCACTGTCAAGTTTCCCGAGTGG--AATATCCTGTGGGGAAGCA 4227  
QY 4417 GGCATCCACGCAACCCCTGGAGTCAGCCACCGCTTTGGCTGTTTTCACACAATGGGGTCTCT 4476  
Db 4228 CGCGGTGCAGACAACTGGAATCAGCCACTGCCCATTTGCTGTCTACAGTGGGGTCTCT 4287  
QY 4477 GTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGCGAGTCAACCACTAGTGG 4536  
Db 4288 GTACATTACTGAAACTGATGAGAAAGAAAATTAACCGGATAAAGCAGGTCAACAACAGATGG 4347  
QY 4537 AGAGATCTCACTCGTTGCTGGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTG 4596  
Db 4348 AGAAATCTCCTTAGTGCGCGGAATACCTTCAGAGTGTGACTGCAAAAATGATGCCAACTG 4407  
QY 4597 TGATTTGTTTTCTGGAGACGATGTTTATGCCAAGGATGCAAAAGTTAAATFACCCCATCTTC 4656  
Db 4408 TGACTGTTACCAGAGTGGAGATGGCTACGCCAAGSATGCCAACTCAGTGCCCATCTC 4467  
QY 4657 CTTGGCTGTGTGCTGATGGGAGCTCTACGTGCCGACCTTGGGAACATCCGGAATTCG 4716  
Db 4468 CTTGGCTGCTTCTCCAGATGGTACACTGTATATTGACAGATCTAGGGAATATCCGGATCCG 4527  
QY 4717 GTTATCCGGAAGAACACAGCCTTTCTCAACACCCCAAGCATGTATGAGCTGTCTTCAAC 4776  
Db 4528 GGCTGTGTCAAAGAAATAAGCCTTTACTTAACTCTATGAACTTCTATGAAGTTGCGTCTCC 4587  
QY 4777 AATTGACCAAGGAGCTCTATCTGTTTGTATACCAACCCGCAAGCACCTGTACACCAAGCCT 4836  
Db 4588 AACTGATCAAGAACTCTACATCTTTGACATCAATGTTACTCACCATAATATGTAAGTTT 4647  
QY 4837 GCCACAGGAGACTACCTGTACAACTTCACTACCTGCGGACCGGACATCAGACTCAT 4896  
Db 4648 AGTCACTGGTGTATACCTTTTACAAATTTTAGCTACAGCAATGACAATGATATGCTGT 4707  
QY 4897 CACAGACAAATGGCAACATGGTAAATGTCCGCCGAGACTCTACTGGGATGCCCTCTG 4956  
Db 4708 GACAGACAGCAATGGCAACACCTTTAGAATTAGACGGGACCCAAATCGCATGCCAGTTCCG 4767  
QY 4957 GCTGTGGTCCAGATGGCAGGTGTACTGTTGGTGACCATGGGCACCAACAGTGCACACTCAA 5016  
Db 4768 AGTGGTGTCTCCTGATAACCAAGTGATATGTTTGAATAAGAAACAAATGGATGTTTGA 4827  
QY 5017 GAGTGTGACACACAAGGACAGAGTTGGCCATGATGACATACCATGGCAATTCGGGCT 5076  
Db 4828 AGGCATGACTGCTCAAGGACTGGAATTAGTTTGTGTTTACTTACCATGGCAATAGTGGCT 4887



Db 7048 TGGATTTTCATGGTGGCTGTATGACCCACTCACAAAATTAATCCACTTTGGAGAAAGAGA 7107  
QY 7294 TTATGATGTGCTGGCGGACGCTGGACTAGCCAGACACACGAGCTGTGGAAGCACCTTAG 7353  
Db 7108 TTATGACATTTTGGCAGGACGCTGGACAACACCTGACATAGAAATCTGGA--AAAGAAT 7164  
QY 7354 TAGCAGCAACGTCATGCTCTTTAAATCTCTATATGTTTCAAAAACAACACCCCATCAGCAA 7413  
Db 7165 TGGGAAGGACCCAGCTCCTTTTAACTTGACATGTTTAGGAATAACAACCCCTGCAAGCAA 7224  
QY 7414 CTCCAGGACATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATT 7473  
Db 7225 AATCCATGACGTGAAGATTACATCACAGATGTTAAACAGCTGGCTGGTGACATTTGGTTT 7284  
QY 7474 CCAGCTACACAACGCTGATCCCTGTTTATCCCAAACACAGACATGGATGCCATGGAACCCCTC 7533  
Db 7285 CCATCTGCACAATGCTATCTCTGATTCCCTGTTCCTCAAAATTTGATTAAACAGAACCTTC 7344  
QY 7534 CTACGAGCTCATCCACACACAGATGAAAACGCGAGAGTGGGACAACACCAAGTCTATCCT 7593  
Db 7345 TTACGA-----ACTTGTGAAGAGTCAGCAGTGGGATGATATACCGGCCCATCTT 7392  
QY 7594 CGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCTTTGTACCTTAGAACGGTTTGA 7653  
Db 7393 CGGAGTCCAGCAGCAAGTGGCGCGGCGAGGCCAAGGCCCTTCTGTGCTGGGGAAGATGGC 7452  
QY 7654 CCAGCTCTATGGTCCACAATCACACAGCTGCCAGCAGGCTCCAAAAGACCAAGAAAGTTTGC 7713  
Db 7453 CGAGGTGCAGGTGAGCCGGCGCGCGCGCGC-----GGCGCGCAGTCTGGCTGTGTTCCG 7509  
QY 7714 ATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGAC 7773  
Db 7510 CACGGTCAAGTCTGATCGGCAAGGGCGCTCATGTGCGCGCTCAGCCAGGGCGCGTGCA 7569  
QY 7774 CACAGACATCATGCTGGCCCAATGAGGATGGGCGAAGGGTTGCTGCCATCTTGAACCA 7833  
Db 7570 GACCAACGTGCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGCGCTGCTCAACAA 7629  
QY 7834 TGCCCACTACCTAGAGAACCTGCACCTTCAACATGATGGGTGGATACCCATTACTTTGT 7893  
Db 7630 CGCCTTCTACCTGGAGAACCTGCACCTTCAACATGAGGGCAAGGACACGACACTACTTCA 7689  
QY 7894 GAAACCAAGGACCTTCAGAAAGGTGACCTGGCCATCCTGGCCCTCAGTGGGGCGGCGAAC 7953  
Db 7690 CAAGACCAACACGCCCGAGAGCGACCTGGSCACGCTGCGTTGACCAAGCGCGCGCAAGGC 7749  
QY 7954 CCTGGAGAAATGGGGTCAACGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGGAC 8013  
Db 7750 GCTGGAGAACGGCATCAACGTGACCGTGTGCGAGTCCACCAACGCTGGTGAACGCGCAGAC 7809  
QY 8014 TAGACGCTACACAGACATCCAGCTCCAGTACGGGCACTGTGCTTGAACACACGCTACGG 8073  
Db 7810 GCGCAGGTTTCGGGACGCTGAGATGACGTTTCGGCGCGCTGGCGCTGACGTCGCGCTACGG 7869  
QY 8074 GACAAAGTTGGATGAGGAGAAGGACGGGTCTTGGAGCTGGCCCGGCAAGAGCCGTGCG 8133  
Db 7870 CATGACCTTGAACGAGGAGAAGGCGCGCATCTTGGAGCAGGCGCGGCAAGGCGCGCTCGC 7929  
QY 8134 CCAAGCGTGGCCCGCAGCAGCAGAGACTGCGGGAAGGGAGGAAGCCCTGCGGGCCTG 8193  
Db 7930 CCGGCGCTGGCGCGCAGCAGCAGCGCGTGCAGCAGCGCGGAGGAGGCGCGCGCTCTG 7989  
QY 8194 GACAGAGGGGAGAAAGCAGAGGTGCTGAGCAAGGGCGGGTGCAAGGCTACGACGGCTT 8253  
Db 7990 GACGAGGGCGAGAAAGCGGAGCTGCTGAGCGCGCGCAAGGTGCAAGGGCTACGACGGGTA 8049  
QY 8254 TTTCTGATCTCTGTGAGAGAGTACCCAGAACTGTTCAGACAGCGCCAAACAACATCCACTT 8313  
Db 8050 CTACGTACTCTCGTGGAGAGTACCCCGAGCTGCGCGCAGCAGCGCCAAACAACATCCAGTT 8109  
QY 8314 CATGAGACAGCGAGATGGGCGCGGAGGTGAC 8345

Db 8110 CCTGCGGCAGACGACATCGGCAGGAGGTAAC 8141  
RESULT 14  
AX250067  
LOCUS AX250067 8689 bp DNA linear PAT 28-SEP-2001  
DEFINITION Sequence 66 from Patent WO0166747.  
ACCESSION AX250067  
VERSION AX250067.1 GI:15864503  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1  
AUTHORS Vernet,C.A., Fernandes,E., Shinkets,R.A., Herrmann,J.L.,  
Majumder,K., Macdougall,J., Mishra,V., Mezes,P.S. and Rastelli,L.  
TITLE Proteins named fctrx and nucleic acids encoding same  
JOURNAL Patent: WO 0166747-A 66 13-SEP-2001;  
Curagen Corporation (US)  
FEATURES  
Location/Qualifiers  
source  
1..8689  
/organism="Rattus norvegicus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10116"  
ORIGIN  
Query Match 40.7%; Score 3397.8; DB 6; Length 8689;  
Best Local Similarity 64.8%; Pred. No. 0;  
Matches 5469; Conservative 0; Mismatches 2712; Indels 254; Gaps 19;  
QY 35 ATGGACGTGAAGGAGAGGAAAGCCTTACCGCTCGCTGACCCGCGCGGACGCCGAGCGC 94  
Db 133 ATGGATGTGAAGGATCGCGACATCGCTCTTTGACCAGGGGACGGTGTGCAAGGAGTGT 192  
QY 95 CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGGCAAG--CCCGCAGAAAATCGTAC 151  
Db 193 CGCTACACCAGCTCCTCTCTGGACAGTGAAGTGTGCGGCTGCGCCACGAGAAGTCTTAC 252  
QY 152 AGCTCCAGCGAGACCTTGAAGGCTTACGACAGGAGCGCCGCTAGCCTATGGCAGCGC 211  
Db 253 AGTTCCAGTGAGACCTTGAAGGCTTATGACCATGACAGCAGAAATGCATATGGAACCGA 312  
QY 212 GTCAAGGACATTTGCGCGAGGAGGCGGAGGAAATTTGCCGACACAGGTGCCAATTCACC 271  
Db 313 GTCACAGACCTGTGTCACCGGAGTCCGATGAGTTTCTAGACAAGGGGCTAATTTTACC 372  
QY 272 CTGCGGGAGCTGGGGCTGGAAGAAAGTAAACGCCCTTACGGGACCTGTACCGGACAGAC 331  
Db 373 CTGGCAGAAATTGGGAATCTGCGAG--CCCTCCCAACACCGAAGTGGTTACTGTTCCGAC 429  
QY 332 ATTGGCTTGCCTTGCCTGAGCACCCTGCTTCCATGGGGCTGGCTCTGATGCCGACATGGAGCT 391  
Db 430 ATGGGATCCTCCACAGGGGCTACTCCCTGAGCACTGGGTCTGATGCGGACTCGGACACC 489  
QY 392 GACACGGTGTGTCCCTGAGCACCCTGCTGCTGTGGGGCGGAGACACCGTCAAGG 451  
Db 490 GAGGAGGGATGTCTCCAGAACATGCCATCAGACTGTGGGACGAGGATAAAATCGAGG 549  
QY 452 GCGAGCTCTGCTGTCAGCGCGGCGCAATTCACACTCACACTCACCGACACCGAGCAT 511  
Db 550 GCGAGCTCTGGCTTGTCCAGCGCGGAGAACTCAGACCTTACTCTGACTGATTTTGACAAT 609  
QY 512 GAAAACACTGAGACTGATCATCCGGCGGCGC---TGCAGAACCAACGCGGCTCCGGAC 567  
Db 610 GAAAATAAATCGGATGACGACAATGGTTCGACCCATTCACCTACATCCTCGTCTAGCCTC 669  
QY 568 GCGCGCGCGCGCTCTGCGCACGCGCCACACCCCGCAACACAGCACCGCGGCTCCATTAA 627  
Db 670 CTCCCATCTGCTCAGCTGCTAGTCTCCCATATCTCCACAGTTAGTCTGCC-AGATGCC 728  
QY 628 CTCCTGAACCGGGGCAACTTCAACGCGGAGGAGCAACCCGCGCGG----- 674

Db 729 ATTGCTAGACAGCAACACCTCCCATCAGATCATGGACACCAACCCCGATGAGGAATTCTC 788  
QY 675 -----CCCCACGGACCACTCGCTCTCCGGAGAGCCCCCTGCCGGCGGCC 722  
Db 789 CCCTAATTATACCTGCTCAGAGCATGCTCAGGGCCCCAGAACGCTCCAGTAGTGGCCC 848  
QY 723 AGGAGCCTGGCCACGCCACGGAGAACTGGCTGCTCAACAGAAACATCCCCCTGGAGACCA 782  
Db 849 TCCGAACCAACACAGCCAGTCAACGCTGAGGCCCTCTGGACCTCTCATAAACACAC 908  
QY 783 GGAACCTAGGAAGCAGCCATTCTTAGGGACATTGCAGGACAACTCATTTAGATGGACA 842  
Db 909 CTTGTCACCAACCACTCCTCTGCCAACTCCCTCAACAGAAACTCACTGACCATAATCGGCG 968  
QY 843 TTCTCGGCGCTCCCGCCA----- 861  
Db 969 GAGTCAATCCACGCCCCAGCTCCTGCACCCCAATGACCTGGCCACCAACGCCGGAGTCCGT 1028  
QY 862 -----TGATGGGCTTACAGTGACGGGCACCTTCT 891  
Db 1029 TCAGCTCCAGGACAGCTGGGTGCTGAACAGTAAACGTGCCGTGGAGACGCGCACTTCT 1088  
QY 892 CTTCAA---GCCTGGAGGCACCTCCCGCTCTTCTGACCAACATCACCAGGTTACCCACT 948  
Db 1089 CTTCAAGACGTCCTCCGGAAGCACACCCCTGTTTCAGCAGCTTCTCCAGGATACCCCTT 1148  
QY 949 GACGTCAGCACAGTGTACTCTCCTCCGCCCCGACCCCTGCCCGCAGCACCTTCGCCCG 1008  
Db 1149 GACCTCAGGACCGTTTATACACCACCCCGCTGCTGCCACGGAATACATTCTTAG 1208  
QY 1009 GCCGGCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAGTGCGCAGCCCTGAG 1068  
Db 1209 GAAGGCTTCAAGCTGAAGAAACCTCCAATACTGCACTGGAAATGCGCCGCCCTGTC 1268  
QY 1069 CGCCATCGTCATCTCAGCCACTCTGGTCATCCTGCTGGGATACTTTGTGGCCATGCACCT 1128  
Db 1269 TGCCATTCCCGTGCCCTCCTCTGGCCATTTTGCTGGCTATTTTCATAGCAATGCATCT 1328  
QY 1129 GTTTGGCCTAAACTGGCACTGAGCCGATGGAGGGGAGATGTATGAGATCAGGAGGA 1188  
Db 1329 GCTCGACTCAATTGGCAACTCCAGCCGGCAGATGGACACACCTTTAACA-----ATGG 1382  
QY 1189 CACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACTGG 1248  
Db 1383 CGTAAGCACCGGCTTACCAGGAACGATGATGTGGCAACAGTGCCATCTGAGGCAAAAGT 1442  
QY 1249 CTTAGACCCCTGCACAGGAAGGCAAGGAACCAAGAAAGCAAGCAAGCCAGTAGTTTCTT 1308  
Db 1443 GCCCTGGTCTGTGAANA----- 1459  
QY 1309 TCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGGAAGCGAGCCTCCAGAA 1368  
Db 1460 -----ACAGCAGCATAGACAGCGCGGAGGCAGAAAGTCGGTCCGCGGTGACACAGGA 1511  
QY 1369 GATTCTCTCGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAA 1428  
Db 1512 AGTCCACACAGGGGTGTTTGGAGGTCCAGATTTCATCAGTCAGCCTCAGTTCTTAA 1571  
QY 1429 ATTCAATGTCTCTGGGAAGGCAGCCCTGGTTGGATTTATGGCAGAAAGCCCTCCC 1488  
Db 1572 GTTCAACATCTCCCTGGGAGGATGCCCTCTTCGGCGTCTACATAAGAAGAGACTGCC 1631  
QY 1489 TCCTTCACATACACAGTTTGTGAGCTGTGAGCTGCTGGAGGAGGCTCCTAACCCCA 1548  
Db 1632 ACCATCTCATGCACAGTATGACTTTCATGGAACGCCCTGGACGAA-----A 1676  
QY 1549 GGAGCGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGGAACCTGTGCCCCCTCCAG 1608  
Db 1677 GGAGAAGTGGAGTGTGGTTCAGTCAACCAGGGAACGCCGGAGCATCCAGACCTGGTGCA 1736  
QY 1609 CCATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTACAA 1668

Db 1737 GAACGAGGCTGTGTTCTGTGCAGTACTTGGATGTGGGCTGTGGCACCTCGCCTTCTACAA 1796  
QY 1669 TGACGGAAAGGAGTCAAGAAGTGGTTTCTTCTCACCACCTGCCATTGAGTCGGTGGATAA 1728  
Db 1797 TGACGGCAAGGACAAGAGATGGTCTCTCTCAATACGGTTGTCTTAGATTCAATGCAGGA 1856  
QY 1729 CTGCCCCAGCAACTGCTATGGCAATGGTGAATGCTGCTCTCTGGGACCTGCCACTGCTTCT 1788  
Db 1857 CTGTCCACGAAACTGCCACGGGAACGGGAATCGGTGTCTGGACTGTGTCACTGTTTCCC 1916  
QY 1789 GGGTTTCTTGGCCCCGACTGTGGGAGAGCCTCTGCCCCGCTGCTGTAGCGGAAATGG 1848  
Db 1917 AGGATTCTTAGTGCAGACTGGCTAAAGCTGCCTGCCCTGTTCTGTGCAGTGGGAATGG 1976  
QY 1849 CCAATACATGAAGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGCTGAGTGCATGT 1908  
Db 1977 ACAGTATTCCAAAGGACATGCCAGTGTCTACAGTGGCTGGAAGGAGAGCAATGCGATGT 2036  
QY 1909 GCCCACCAACAGTGTATCGATGTGGCTGCAGCAACCATGGCACCTGCATCACGGGCAC 1968  
Db 2037 GCCCATGAACAGTGCATCGATCCTTCTGTGGGGCCACGGCTCCTGCATTTGATGGAA 2096  
QY 1969 CTGCATCTGCAACCTGGCTACAAGGGGAGAGCTGTGAGGAAGTGGACTGCATGGACCC 2028  
Db 2097 CTGCGTGTGTGAGCTGGCTACAAAGGCGAGCACTGCCAAGAAAGTGGATTGCTTGGATCC 2156  
QY 2029 CACATGTTCAAGGCCGGGTGTCTGCTGAAGAGGCGGAATGCCATTGCTTTGTGGATGGG 2088  
Db 2157 AACCTGTCTCAGCCCATGGTGTCTGTGAACGGAGAGTGTCTATGCAGCCCCGGCTGGG 2216  
QY 2089 AGGCACCACTGCGAGACCCCCAGGSCCACATGCTTAGACCAGTGTTCAGGCCACCGGAAC 2148  
Db 2217 CGGGCTCAACTGCGAGCTGGCGAGGCTCCAGTCCAGTCCACAGACCTGATGGCATGGCAC 2276  
QY 2149 CTTCTCCCGACACCGGGCTTTGCACTGTGACCCAAAGCTGGAATGGACACGACTGTTTC 2208  
Db 2277 TTACCTCCCTGACTCTGGCCTCTGCAACTGTGATCCGAATTGGATGGTCCCAGCTGCTC 2336  
QY 2209 TATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCG 2268  
Db 2337 TGTGAAGTGTCTCAGTAGACTGTGGCACTCACGGCGTCTGCATCGGGGAGCCTGCCG 2396  
QY 2269 CTGCGAGGATGGTGGATGGGGGACCTGCCAGCCAGCGGGCTGCCACCCGCTGTGC 2328  
Db 2397 CTGTGAAGAGGGCTGGACAGGGCGGCTTGTGACCAAGCGGCTGTGCCACCCCGCTGCAT 2456  
QY 2329 CGAGCATGGGACCTGCCCGCAGCGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACA 2388  
Db 2457 TGAGCACGGGACCTGTAAAGATGGCAAAATGTGAATGCCAGAGGGCTGGAATGGTGAACA 2516  
QY 2389 CTGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTCCTGGTTGTGCAA 2448  
Db 2517 CTGCACCATTGAT-----GGCTGCCCTGATTGTGCAA 2549  
QY 2449 TGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGGCTGTGCCAGCTGGGCTG 2508  
Db 2550 CGGTAACGGGAGATGCACACTGGGTCAAGAACAGCTGGCAGTGTCTGCCAGACCGGCTG 2609  
QY 2509 GAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCCCTGGGTGACAGCAAGACAATGA 2568  
Db 2610 GAGAGGGCCGGATGCAACGTTGCCATGGAAACCTCCTGGCTGATAACAAGGATAATGA 2669  
QY 2569 TGGAGATGGCTGGTGGACTGCATGACCCCTGACTGTGCTGCCAGCCCTGTGCCATAT 2628  
Db 2670 GGGAGATGGCTGGTGGACTGCCTGGACCCCTGACTGTGCTGCCCTCAGTCAGCCTGTGAGAA 2729  
QY 2629 CAACCCGCTGTGCCCTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCC 2688  
Db 2730 CAGCCTGCTGTGCGGGGTCTCGGAGCCCTTGGACATCATTCAGCAAGGCCAGACAGAGA 2789  
QY 2689 TGTGTACAGCAGAACTTACACTCTTCTATGACCCGCTCAAGTTCTCTGTTGGGCGAGGA 2748  
Db 2790 CTGGCCTGCGGTGA-----AGTCCTTCTATGATCGTATCAAGCTCTTGGCAGGCAAGGA 2843



Db 5001 CGGGAATTCCTAAAGATCCGCGGGACAGCAGTGGCATGCCCGACACCTGCTCATGCC 5060  
QY 4969 AGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGACCAAC 5028  
Db 5061 TGATAATCAGATCATCACCTTACGGTGGGCACCAACGGAGGCTCAAAGCCGTGTCAAC 5120  
QY 5029 ACAAGGACACAGAGTTGGCCATGATGACATACCATTGGCAATTCCGGCCCTTCTGGCAACCAA 5088  
Db 5121 GCAGAACCTGGAGCTGGCCCTCATGACTTATGATGGAAACACTGGACTCCTAGCCACCAA 5180  
QY 5089 AAGCAATGAAAACGGATGGACAACATTTTATGAGTACGACAGCTTTGGCCCGCTGACAAA 5148  
Db 5181 GAGCGATGAAAACGGATGGACAACCTTTTATGACTATGACCAAGAGGGCCGTCTGACCAA 5240  
QY 5149 TGTGACCTTCCCTACTTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCA 5208  
Db 5241 TGTGACTCGCCCCACGGGGTGGTGACCAAGCCTGCAACGGGAAATGGAGAAATCCATCAC 5300  
QY 5209 TGTCCAGGTAGAGACCTCCAGCAAGGAT---GATGTCACCAATAACCAACCTGTCTGC 5265  
Db 5301 CGTTGACATGAGAACTCCAACCCGTGATAACGATGTCACTGTGATTACCAACCTCTCTTC 5360  
QY 5266 CTCAGGGCCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGC 5325  
Db 5361 AGTGGAGCCCTCTACACCGTGGTACAAGATCAAGTCCGGAACAGTACCAGCTCTGCAG 5420  
QY 5326 CGATGGCTCCTTGGCGTGTCTGTGGCCAAACGGCATGGAGTGGCGCTGCAGACTGAGCC 5385  
Db 5421 CAACGGGACCTGCGCGTCACTGTACGCCAAACGGCATGGCGCTCAGCTTCCACAGCGAGCC 5480  
QY 5386 CCACCTTGTGGCTGGCACCGTCAACCCCAACCGTGGCAAGAGGAATGTCAACGCTGCCCAT 5445  
Db 5481 CCACGTCCTCGCAGGCACCCCTCACCCCAACCATCGGGCGCTGTAAACATCTCCCTGCCCAT 5540  
QY 5446 CGACAAGCCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGT 5505  
Db 5541 GGAGAACGGCTGAACCTCCATCGAGTGGCGCTGAGGAAGGAACAGATTAAAGGCAAGT 5600  
QY 5506 CACTGTCTTTGGCGCGCGTGGGGTGACAAACCGAATCTCCTATCTCTGGACTTTGA 5565  
Db 5601 CACCATCTTTGGAGGAAGCTTCGGTCCACGGAAGGAACCTCCTGTCCATTGATTATGA 5660  
QY 5566 TCGCGTAACACGCACAGAGAAGATCTATGATGACCACCGCAAGTTCACCCCTTCGGATTCT 5625  
Db 5661 CCGAATATCCGCACCTGAGAAGATCTATGACGACCAACCGAAGTTCACCCCTGAGATCAT 5720  
QY 5626 GTACGACAGCGGGCGGCCAGCCCTCTGGTCACCCAGCAGCAGGCTGAATGGTGTCAA 5685  
Db 5721 TTATGACAGGTGGCGCGCCCTCTCTGTGGCTCCCAAGAGTGGACTGGCGGCGCTCAA 5780  
QY 5686 CGTGACATACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCGTGAAAG 5745  
Db 5781 TGTCTCTACTTCTTCAAACGGCGCCTGGCCGGCTCCAGCGCGGGCCATGAGCGAGAG 5840  
QY 5746 AATGGAATACGACGAGCGCGGCATCAATCCAGGATCTTCGCTGATGGGAAGACATG 5805  
Db 5841 GACAGACATTGACAAGCAAGGCGGATTGTGTCCCGAATGTTCGCCGACGGGAAAGTCTG 5900  
QY 5806 GAGCTACACATACTTAGAGAAGTCCATGGTGTCTGCTACTACACAGCCAGAGGAGGATAT 5865  
Db 5901 GAGCTATTCCTACCTTGACAAGTCCATGGTCTCTCTGCTGCAGAGCCAGCGTCAGTACAT 5960  
QY 5866 CTTTGAGTTCGACAAGAAATGACCGCCTCTCTTCTGTGACGATGCCCAACGTGGCGCGGCA 5925  
Db 5961 ATTTGAATATGACTCCTCTGACCGCCTCCACGCAAGTCCATGCCAGTGTCTCGCCCGGCA 6020  
QY 5926 GACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGG 5985  
Db 6021 CAGCATGTCCACGCACACCTCCATTGGCTACATCCGGAACATTTACAACCCACGGGAAG 6080  
QY 5986 CAATGCTCAGTCAATACAGGACTTCACTGAGGATGGGACCTCTTTCACACCTTCTACCT 6045

Db 6081 CAACGCCCTCGGTCACTCTTTGACTACAGTATGACGGCGCATCTCTGAGACGCTCTTTCCT 6140  
QY 6046 GGGCACTGGCCGAGGTTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCT 6105  
Db 6141 GGGCACCGGGCGCAGGTGTTCTATAAGTACGGAATACTGTCCAAGTTATCGGAGATCGT 6200  
QY 6106 CTATGACACCAACCAAGGTCAGTTTTCACCTATGACGAGACGGCAGGCATGCTGAAGACCAT 6165  
Db 6201 CTACGACAGCACTGCCGTCACTTTCGGCTATGACGAGACCACTGGCGCTCTGAAGATGGT 6260  
QY 6166 CAACCTACAGAAATGAGGGCTTCACCTGCAACCATCCGCTACCGTCAAGTTGGGCCCTGAT 6225  
Db 6261 GAATCTCCAAAGCGGGGCTTCTCCTGTACCATCAGGTACCGAAAGTTCGGGCCCTCGT 6320  
QY 6226 TGACCGACAGATCTTCGGCTTCACCTGAGGAAGGCATGGTCAACGCCCGTTTGTACTACA- 6284  
Db 6321 GGACAAGCAGATTACAGGTTCTCTGAGGAAGGCATGATCAACGCCAGGTTTCGATTACAC 6380  
QY 6285 --ACTATGACAAACAGCTTCCGGGTGACCAAGCATGCAGGCTGTGATCAACGAGACCCCACT 6342  
Db 6381 CTACCACGACAAACAGCTTCCGCATCGCCAGCATCAAGCCCGTCATCAGTGAGACTCCCT 6440  
QY 6343 GCCCATTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGACAGTTTGGGAAGTT 6402  
Db 6441 TCCCGTTGACCTCTACCGCTACGATGAGATTCTGGCAAGGTGGAACACTTCGGCAAGTT 6500  
QY 6403 TGGTGTCAATTACTATGACATTAAACAGATCATCACACAGCTGTGATGACCCACACCAA 6462  
Db 6501 CGGGGTCACTACTACGACATCAACCAAGATCATCACCTGCCGTGATGACACTCAGCAA 6560  
QY 6463 GCATTTTGTGATGATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCGCTCGTCTCAT 6522  
Db 6561 GCATTTTGACACCCATGGCGCATCAAGGAAGTGCAGTATGAGATGTTCCGGTCCCTCAT 6620  
QY 6523 GTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGGAGCTGAAGGT 6582  
Db 6621 GTACTGGATGACGGTGCAATATGACAGTATGGCAGGTGATCAAGAGGGAACCTGAAACT 6680  
QY 6583 AGGACCCCTACGCCAAATACCACCTCGCTACTCCTATGATGATGATGACGACCTCAATGGGAACCT 6642  
Db 6681 GGGGCCCTATGCCAACACCAACCAAGTACACCTATGACTACGACGGGACGGCCAGCTCCA 6740  
QY 6643 GACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCT 6702  
Db 6741 GAGTGTGGCGCTCAATGACCGGCCCTACCTGGCGTTATAGCTATGACCTCAATGGGAACCT 6800  
QY 6703 GCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCCGCA 6762  
Db 6801 GCACCTGTATAACCCAGGAACAGTGTCTCGCTCATGCGGTACGCTATGACCTCCGTGA 6860  
QY 6763 CCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGATGGCTTCTCTGAGGCA 6822  
Db 6861 CCGGATAACCAAGCTAGGGACGTGCAGTACAAAATCGATGATGATGGCTATTTATGCCA 6920  
QY 6823 GCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATCAAGGCCTACAACCG 6882  
Db 6921 GAGAGGATCTGACATCTTTGAATACAACTCCAAGGCCCTTCTAACGAGAGCGTACAACAA 6980  
QY 6883 GGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCTGGGGCGCGCGCTGTCCAGCAA 6942  
Db 6981 GGCCAGCGGGTGGAGTGTGCAGTACCGCTATGATGGCGTGAGCGCCGGCTTCTCTACAA 7040  
QY 6943 GAGCAGCCACAGCCACCACTGAGTTCTTCTATGACAGACCTGACCAACCCCAAGGT 7002  
Db 7041 GACCAACCTGGGCCACCACTACAGTACTTCTATTCCGACCTCCACCAACCCCAACGTAT 7100  
QY 7003 CACCCACCTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGCAGG 7062  
Db 7101 CACCCATGTTTACAACCACTCCAACCTGAGATCACTCACTCTACTATGACCTCCAGGG 7160  
QY 7063 ACACCTCTTTGCCATGGAGCTGAGCAGTGGTGTGATGAGTTTACATAGCTTGTGACACAT 7122  
Db 7161 CCACCTCTTTGCCATGGAGAGCAGTAGTGGGAAGAGTACTATGTTGCCTCAGATAACAC 7220

Qy 7123 CGGGACCCCTCTTCTGCTCTTTAGTGGAACAGGTTTGATGATCAAGCAAAATCCTGTACAC 7182

Db 7221 CGGGACTCCTCTGGCTGTTTTAGTATCAATGGCCTCATGATCAAGCAAAATCCTGTACAC 7280

Qy 7183 AGCCTATGGGAGATCTACATGGATACCAACCCCAACTTTCAGATCATATAGGCTACCA 7242

Db 7281 AGCCTATGGGAGATTTACTATGACTCCAATCCAGACTTTCAGATGTTTATCGGCTTCCA 7340

Qy 7243 TGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGCGGAGATTATGATGT 7302

Db 7341 CGGAGGCCTCTACGACCCCTCACCAAGCTCGTTCACTTTACGAGCGGTGATTATGACGT 7400

Qy 7303 GCTGGCCGAGCGTGGACTAGCCAGACCAAGCTGTGGAAGCACTTAGTAGCAGCAA 7362

Db 7401 GCTGGCAGGACGGTGGACGTCCCCGACTACACCATGTGGAGAA---TGTGGGCAAGGA 7457

Qy 7363 CGTCATGCCCTTTAATCTCTATATGTTCAAAAACAAACCCCATCAACAACTCCCAGGA 7422

Db 7458 GCCAGCCCCCTTCAACCTGTACATGTTCAAGAAACAAATCCACTCAATGAGCTGGA 7517

Qy 7423 CATCAAGTGTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACA 7482

Db 7518 TTTAAAGAACTACGTGACAGACGTGAAGAGCTGGCTCGTGATGTTTGGATTTCAGCTCAG 7577

Qy 7483 CAACGTGATCCCTGGTTATCCCAACGAGGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACA 7542

Db 7578 CAACATCATCTCTGGATTCCCAAGAGCCAAATGATTTTGTGCTCCCTCCCTATGA--- 7634

Qy 7543 CATCCACACAGATGAAAACGAGGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACA 7602

Db 7635 -----ACTGTGAGAGAGCAAGCAAGTGAGAATGGACAGCTCAATACAGGTGTCCA 7685

Qy 7603 GTGTGAAGTACAGAAGCGAGCTCAAGGCTTTGTACCTTAGAACGGTTTGACCAAGCTCTA 7662

Db 7686 GCAGACAACAGAGAGGCAATAACAGGCTTTCTGCTCTAGAGGACAGGTCTCTCTAA 7745

Qy 7663 TGGCTCCACAATCACAGTGCAGGCTCCAAAGACCAAGAAAGTTGTCATCCAGCGG 7722

Db 7746 AAAGCTCCATGCAGGCTCCGAGAGAAAGC-----AGGCCACTGGTTGTACGACCAC 7799

Qy 7723 CTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCAACAGACAT 7782

Db 7800 GCCCATCATCGGCAAGGCATCATGTTCCGCTTCAAGAGAGGGCGGTGACCAACAGCGGT 7859

Qy 7783 CATCAGTGTGGCCAATGAGATGGGCGAAGGGTTCTGCTGCCATCTTGAACCATGCCCACTA 7842

Db 7860 GTCTAGCATCGCCAGTGAGACAGCCGCAAGGTAGCATCCGTGTTGAACAACGCCTACTA 7919

Qy 7843 CTTAGAGAACTGCACTTCAACATGATGGGGTGGATACCCATTACTTTGTGAACCCAGG 7902

Db 7920 CTTGGAAGAAGATGCATACAGCATCGAGGCAAGGACACACACTACTTCTGTTGAAGATCGG 7979

Qy 7903 ACCTTCAGAAGGTGACCTGGCCATCCTGGGCCTCAGTGGGGGGCGGCAACCTTGGAGAA 7962

Db 7980 TGCAAGCGGACGGTGACCTGGTTACGCTGGGACCAACCATTTGGCGCAAGGTCTGGAGAG 8039

Qy 7963 TGGGTCAAACGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTA 8022

Db 8040 CGGGGTGAACGTGACCGTGTACAGCCACAGCCACGCTGCTGGTGAACGCGCAGGAGTGAAGGTT 8099

Qy 8023 CACAGACATCCAGCTCCAGTACGGGGGCACTGTGCTTGAACACACAGCTACGG----- 8073

Db 8100 CACCAACATTTGAATTCAGTACTCCACGCTGCTGCTCAGCATACTGCTACGGGCTCACCCC 8159

Qy 8074 GACAAGCTTGGATGAGGAGAGGCAAGGCTCCTGGAGCTGGCCCGGACAGAGCCGTGCG 8133

Db 8160 CGACACACTGGATGAAGAGAGGCGCGCGTCTCTGGACCAAGCGGACAGAGGCGCTGGG 8219

Qy 8134 CCAAGCTGGGCGCGGAGCAGCAGAGACTGCGGGAAGGGGAGGAAGGCTTGGGGGCTG 8193

Db 8220 TACTGCTGGGCCAAGGAGCAGAGAAAGCCAGGACGAGAGAGGCGGAGAGGCGGCTGTGTG 8279

Qy 8194 GACAGAGGGGAGAGCAGAGGCTGCTGACACAGCGGGGTGCAAGGCTACGACGGCTT 8253

Db 8280 GACGGAGGGCGAGAGCAGCAACTCCTGAGCAGCGGACGGGTGCAAGGTTATGAGGGCTA 8339

Qy 8254 TTTTCGTGATCTCTGTCGAGCAGTACCCAGAACTGTTCAGACAGCGGCCCAACATCCACTT 8313

Db 8340 TTACGTGCTTCCGGTGGACAGTACCCAGAGCTGGCAGACAGTAGCAGCAACATCCAGTT 8399

Qy 8314 CATGAGACAGAGCAGATGGCCGCGGAGGTGACAGA 8348

Db 8400 CTTAAGACAGATGAGATGGGAAAGAGGTAAACAA 8434

RESULT 15

AF086607

LOCUS

DEFINITION Rattus norvegicus neurestin alpha mRNA, complete cds.

ACCESSION AF086607

VERSION AF086607.1 GI:5712200

KEYWORDS

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 8689)

AUTHORS Otaki, J.M. and Firestein, S.

TITLE Neurestin: putative transmembrane molecule implicated in neuronal development

JOURNAL Dev. Biol. 212 (1), 165-181 (1999)

MEDLINE 93350226

PUBMED 10419693

REFERENCE 2 (bases 1 to 8689)

AUTHORS Otaki, J.M.

TITLE Direct Submission

JOURNAL Submitted (25-AUG-1998) Department of Biological Sciences, Columbia University, Fairchild Center, New York, NY 10027, USA

FEATURES

source Location/Qualifiers

1..8689

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague Dawley"

/db\_xref="taxon:10116"

/tissue\_type="olfactory bulb"

133..8430

/note="putative neuregulin-like receptor; proposed as a receptor for odorant receptors; may function for target recognition and synapse formation; similar to Homo Sapiens gamma heregulin, Drosophila Ten(a), mouse DOC4, and Drosophila Odz/Ten(m); three additional variants (beta, gamma, and delta) identified"

/codon\_start=1

/product="neurestin alpha"

/protein\_id="AAD47383.1"

/db\_xref="GI:5712201"

/translation="MDVKDRHRSLTFRGCKECRYTSSSLDSHEDCRVPTQKYSSE TLKAYDHSRMHYGNRVTDLVHRESDFSRQGANFTLAELGLICEFPSPHRSGYCSMDGE LHQYSLSTGSDADSDTEGMSPEHAIRLWGRGKISRRSSGLSSRENSALTITDSNENKSDDDNGRPIPTSSSSLLPSAQLPSSHNPPVSCQPLDNTSHQIMDTPDEEF SPNSYLLRACSGPQQAASSGPPNHHQSQTLPPLPPPHNHTLSHHSSANSLNRLNRRSQIHAPAPNDLATTPEVQLQDSVNLNSNVPLETRHFLFKTSSTPLFSSSPGYLTSVTVPPLRLPRNTFSRKAFLKPKSKYCSWKCAALSAIAAALLAILLA YFIAMHLGLNWQLQPADGHTFNNGVTRTGLPGNDDVATVPSSGKVPWSLKNSSIDSGE AEVGRRTQEPVPGVFRSQIHI SQPQLKFNISLGKDALFGVYIRRLGPPSHAQYDF MERLDGKEKSVSVESPRRISQITLVQNEAVFVQLDVLGHLAFYNDGDKEMVSFN TVLDSVQDCPRNCHNGECVSGLCHCFPGFLGADCAKACPVLCSGNGQYSGKTCCQ YSGWKGAECIDPSCGGHSCIDGNCVCAAGYKGEHCBEVDCLDPTCSSHGVCVNGECLSPGWGLNCELARVQCPDQCSGHGTYLPDGLCNDPWNMGPDSCSVEVCS VDCGTHGVCIIGACRCCEGTGAACDORVCHPRCIEHGTCKDGKCEGREGNWEHCTI DGCPDLNCGNRCITLGQNSWCVCQVQGWGPGCNVAMETSCADNKDNEGDLVDCLDP DCCLQSACQNSILCRGSRDPLDIIQQGTDWPAVKFYDRIKLAKGDSHTIIPGDNPFNSLSVSLIRGQVVTDTGTPLVGVNVSFVKYPKYGYTITRQDGTFDLIANGGSALTILH FERAPFMSRERTVWPPWNSFYAMDTLVMKTEENSIPSCDLSGFRPDPPIIISPLSTF

CDS

FSASPAANPIVPETQVLHHEIELPGTNVKLRYLSRTAGYKSLKIKITMTQSTVPLNLI  
RVHLMVAVEGHLFQKSFQASPNLAYTFIWDKTDAYQQRVYGLSDAVSVGEYETCP  
LILWEKRTALLOGFELDPNLSLGNLWLDKHTLNVKSGILLKGTGENQFLTQQPAIITS  
IMGNRRRSISCPSCNGLAENKLLAPVALAGIDSLFVGFDFNYIRRIFFPSRNVTSI  
LELNKKEFKHSNPGKXYILAVDPVTSGLYVSDNRSRRIYRVKSLSGADKLAGNSEVV  
AGTBEQCLPFDEARCGGKAVDATLMSPRGIAVDKNGMLFVDAATMRKVDQNGIIS  
TLILGNDLTAVRPLSCDSSMDVAQVRLEWPTDLAVNPMNLSLYVLENNVILRITENHQ  
VSIAGRPMHCQVPGIDYSLKLAHSALESASATAISHTGLVYITETDEKINRLRQ  
VTNGEICLLAGAAASDCCKNDVNCICYSGDDAVATAILNSPSSLAVAPDGLIYIAD  
LGNIRIRAVSKNKPVLNAFQYEAASPEQELVFNADGIHQYTVSLVTEGTYNFTY  
SADNDVTELDINNNSLKIIRDSSGMPRHLLMPNQIITLTVTNGGLKAVSTQNLEL  
GLMTYDNGTGLLATKSDETGWTTFYDYDHEGRLNTRPTGVVTSLSHREMEKSIITVDI  
ENSNRDNVTITNLNSVEASVTVVQDQVRNSYQLCSNGTLRVMYANGMGSVFSHSEPH  
VLAGTLTPTIGRCNISLPMENGLNSIEWRLRKEQIKGKVTIFGRKLRVHGRNLLSIDY  
DRNIRTEKIYDDHRKFTLRIIYDQVRPFLWPSGLAAVNSYFFNGELLAGLQRGAM  
SERTDIDKQRIVSRMFADGKVMYSYLDKSMVLLQLSQORQYIFEYDSSDRLHAVTMP  
SVARHSMSTHSIGYIRNIYNPPESNASVIFDYSDDGRIILKTSFLGTGQVFFKYGKL  
SKLSEIIVYDSTAVTFGYDETGVLMKNVLQSGGFSTIRYKVGPLVDKQIYRFSEEG  
MINARFDYTYHDNSFRITASIKPVISETPLPVDLYRYDEISGKVEHFGKGVIIYDINQ  
IITTAVMTLSKHFDTHGRIVEQYEMERSLMYWMTVQYDSMGRVIKRELKLGPYANTT  
KYTYDYDGDGQLQSVAVNDRPTWYSYDLNGLNLHLPNGNSARLMPRLYDLRDRITRL  
GDVQYKIDDDGYLCQSGDIFEYNSKGLLTRYANKASGWSVQYRYDGVSPRASXYTNL  
GHLQYFYSDLHHPTRITHYNHNSSEITSLYDQLQHLFAMESSEEEYVVASDNTG  
TPLAVFSINGLMIKQLQYATYGEIYYDSNPDFQMVIFGHGGLYDPLTKLVHFTQDYD  
VLAGRWTSPDYTMWRNVGKEPAPFNLYMFKNNPLSNELDKNYVTVDKSWLWFMFGFQ  
LSNIIIPGPRAKMYFVPPPYELSESOASENGQLITGVQOQTERHNOAFLEAGQVISK  
KLHAGIREKAGHWFATTPIIGKIMFAIKEGRVTTGVSSIASSEDSRKVASVLNNAYY  
LDKMHSYIEGKDYHFKVIGAADGDLVTLGTTIGRKVLESQVNVTVSQPTLLVNGRTR  
RFTNIEFQYSTLLSIRYGLTPTDLEEKARVLDQARQALGTAWAKEQOKARDGREG  
SRLWTEGEKQOLLSTGRVQYEGYVLPVEQYPELADSSSNIQFLRQNEWNGKR"

ORIGIN

Query Match		40.7%;	Score 3397.8;	DB 10;	Length 8689;
Best Local Similarity		64.8%;	Pred. No. 0;		
Matches 5469;		Conservative	0;	Mismatches 2712;	Indels 254; Gaps 19;
QY	35	ATGGACGTGAAGGAGAGAGAGCCCTTACCGCTCGCTGACCCCGCGCGGACGCCGACGCG	94		
Db	133	ATGGATGGAAGGATCGCGGACATCGCTCTTTGACCAGGGGACGGTGTGGCAAGGAGTGT	192		
QY	95	CGCTACACCGAGCTCGTCCGCGGACAGCGAGGAGGGCAAAG---	151		
Db	193	CGCTACACCGAGCTCCTCTCTGGACAGTGAGGACTGCCGTGTGCCACGCAAGATCCTAC	252		
QY	152	AGCTCCAGCGAGACCCCTGAAGGCCTACGACCGAGGACGCCCGCTAGCCTATGGCAGCCG	211		
Db	253	AGTTCAGTGAGACCCCTGAAGGCTTATGACCATGACAGCAGATGCACATGGAACCCGA	312		
QY	212	GTCAAGGACATGTGCCGAGGAGCGCGAGGAATCTTGCGGACAGGTGCCAATTCACC	271		
Db	313	GTCACAGACCTGGTGACCCGGGAGTCCGATGAGTTTCTAGACAAGGGGCTAATTTAC	372		
QY	272	CTGCGGGAGTGGGGCTGGAAGAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGAC	331		
Db	373	CTGGCAGAATGGGAATCTGCGAG---CCCTCCCCACACCGAAGTGTTACTGTTCGAC	429		
QY	332	ATTGGCCTGCCCCAATGCGGCTACTCCATGGGGCTGGCTGTGATCCGACATGGAGGCT	391		
Db	430	ATGGGGATCCTCCACCGAGGCTACTCCCTGAGCACTGGGTCTGATGCGGACTCGGACACC	489		
QY	392	GACACGGTGTCTCCCTGAGCACCCCGTGGCTGTGGGGCGGAGCACACGGTCAAGG	451		
Db	490	GAGGGAGGGATGTCTCCAGAACATGCCATCAGACTGTGGGGACGAGGATAAAATCGAGG	549		
QY	452	CGCAGCTCTGCTGTCCAGCCCGSGSCAAATTCCAATCTCACACTCACCGACACCGAGCAT	511		
Db	550	CGCAGCTCTGGCTTGTCCAGCCCGGAGAACTCAGCCCTTACTCTGACTGATTTCTGACAAT	609		
QY	512	GAAAACACTGAGACTGATCATCCGGGGGGCC---	567		
Db	610	GAAAATAAATCGGATGACGACAATGGTCGACCCCATTCACACCTACATCCTCGTCTAGCCTC	669		
QY	568	GCCGCCGCCCGCGCTCTCGCACCGCCACACCCCCAACCCAGCACCGCGGCTCCTCAATAA	627		

Db	670	CTCCCATCTGCTCAGCTGCCTAGCTCCCATATAATCCTCCACCAGTTAGCTGCC-AGATGCC	728		
QY	628	CTCCCTGAACCGGGCAACTTCACGCCGAGGAGCAACCCCGAGCCGG-----	674		
Db	729	ATTGCTAGACAGCAACACCTCCCATCAGATCATGGACACCAACCCCGATGAGGAATTCTC	788		
QY	675	-----CCCCACGGACCACTCGCTCTCCGGAGAGCCCCCTGCCGGCGGCGCCC	722		
Db	789	CCCTAATTCTATACCTGCTCAGAGCATGCTCAGGGCCCCCAGCAAGCCTCCAGTAGTGCCCC	848		
QY	723	AGGAGCCTGCCACGCCCCAGGAGAACTGCTGCTCAACAGCAACATCCCCCTGGAGACCA	782		
Db	849	TCCGAACCAACACAGCCAGTCAACGCTGAGGCCCTCTGCCCACCTCTCATACCACAC	908		
QY	783	GGAACCTAGGCAAGCAGCCATTCTCTAGGACATTCAGGACCAACCTCATTTGAGATGGACA	842		
Db	909	CCTGTCCCACCACTCCTCTCTGCCAACTCCCTCAACAGAAACTCACTGACCAATCGGCG	968		
QY	843	TTCTCGGCGCCTCCCGCCA-----	861		
Db	969	GAGTCAAATCCAGCCCCAGCTCCTGCACCCAATGACCTGGCCACCAACGCGGAGTCCGT	1028		
QY	862	-----TGATGGGCTTTACAGTGACGGGCACTTCCT	891		
Db	1029	TCAGCTCCAGACAGCTGGTGTGAACAGTAACGTCCCGCTGGAGACGCGGCACTTCCT	1088		
QY	892	CTTCAA--GCCTGGAGGACCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACT	948		
Db	1089	CTTCAAGACGTCTCCGGAAGCACACCCCTGTTTTCAGAGCTCTTCTCCAGGATACCCCTT	1148		
QY	949	GACGTCCAGCACAGTGTACTCTCCTCCGCCCCGACCCCTGCCCCGAGCACCTTCGCCCC	1008		
Db	1149	GACCTCAGGGACCGTTTATACACCACCCCGCCTGCTGCCACGGAATACATTCTCTAG	1208		
QY	1009	GCCGGCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAG	1068		
Db	1209	GAAGGCTTCAAGCTGAAGAAACCTCCAAATACTGCAGTTGGAATGCGCCGCTGTC	1268		
QY	1069	CGCCATCGTCATCTCAGCCACTCTGTGTCATCCTGTGGCATACTTTTGGCCATGCACCT	1128		
Db	1269	TGCCATTGCGGCTGCCCTCCTTCTTGCCCATTTTGTGGCCTATTTTCATAGCAATGCATCT	1328		
QY	1129	GTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGGAGGA	1188		
Db	1329	GCTCGGACTCAATTGGCAACTCCAGCCGCGAGATGGACACACCTTTAACA-----ATGG	1382		
QY	1189	CACAGCCAGAGTTGGCCTGTGCCAACCGACAGTCTCCCTATACCCCTCAGGGGGCACTGG	1248		
Db	1383	CGTAAGGACCGGCTTACCAGGAAACGATGTGGCAACAGTGCCATCTGGAGGCAAGT	1442		
QY	1249	CTTAGAGACCCCTGACAGGAAAGGCAAGGAACCAACAGAAGGAAGCCAGTAGTTTCTT	1308		
Db	1443	GCCCTGGTCGTTGAAA-----	1459		
QY	1309	TCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGGAAGCGGAGCCTCCCAGAA	1368		
Db	1460	-----ACAGCAGCATAGACAGCGCGAGGCAGAAAGTCGGTCCAGCGGTGACACAGGA	1511		
QY	1369	GATTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAA	1428		
Db	1512	AGTCCCAACCAAGGGGTGTTTGGAGGTCCAGATTTCATTCAGTTCAGCTTCTTTAAA	1571		
QY	1429	ATTCAATGTGTCTCTGGAAAGGACGCCCTGGTTGGCATTTTATGGAGAAAAGCCCTCCC	1488		
Db	1572	GTTCAACATCTCCCTGGGAAGGATGCCCTCTTCGGCGCTCTACATAAGAAGAGGACTGCC	1631		
QY	1489	TCCTTACATACACAGTTTGACTTTTGTGGAGCTGTCTGGATGGCAGGAGGCTCTTAACCCA	1548		
Db	1632	ACCATCTCATGCACAGTATGACTTTCATGGAACGCTTGACGGAA-----A	1676		
QY	1549	GGAGGCGGAGGCTTAGAGGGGACCCCGGCCAGTCTCGGGGAAGTGTGCCCCCTCCAG	1608		

Db 1677 GGAGAAAGTGGAGTGTGGTCGAGTCAACCCAGGGAAACGCCGGAGCATCCAGACCCCTGGTGCA 1736  
Qy 1609 CCATGAGACAGGCTTCATCCAGTATTTGGATTACAGGAATCTGGCACTTGGCTTTTACAA 1668  
Db 1737 GAACGAGGCTGTGTTCTGTCAGTACTTGGATGGGCCCTGTGGCACCTCGCCTTCTACAA 1796  
Qy 1669 TGACGGAAAGGAGTCAGAAAGTGGTTTCCCTTTCTCACCACCTGCCATTGAGTCGGTGATAA 1728  
Db 1797 TGACGGCAAGGACAAGGAGATGGTCTCCTTCAATACGGTTGTCTTAGATTGAGTCAGGA 1856  
Qy 1729 CTGCCCCAGCAACTGCTATGGCAATGGTGACTGCACTCTCTGGGACCTGCCACTGCTTCCT 1788  
Db 1857 CTGTCCACGAAACTGCCACGGGAACGGCGAATGGGTGTCTGGACTGTGTCACTGTTTCCC 1916  
Qy 1789 GGGTTTCTGGGCCCGGACTGTGGCAGAGCCTCTGTCCCCCGTGTCTGTAGCGGAAATGG 1848  
Db 1917 AGGATTCCTAGGTGCAGACTCGCTAAAGCTGCTGCCCTGTTCTGTGCGAGTGGGAATGG 1976  
Qy 1849 CCAATACATGAAAGGAGATGCTTGTGCCACAGTGGCTGGAAGGCCGTGAGTGCATGT 1908  
Db 1977 ACAGTATTCCAAAGGACATGCCAGTGTCTACAGTGGCTGGAAGGAGCAGAAATGCGATGT 2036  
Qy 1909 GCCACCAACAGTGTATCGATGTGGCCTGACAGAACCATGGCACCTGCATCACGGGCAC 1968  
Db 2037 GCCATGAACCAAGTGCATCGATCCTTCTGTGGGGCCACGGCTCCTGCAATGATGGAA 2096  
Qy 1969 CTGCATCTGCAACCCCTGGCTACAAGGGCGAGAGTGTGAGGAAGTGGACTGCATGGACCC 2028  
Db 2097 CTGGTGTGTGACGCTGGCTTACAAAGGGCGAGCACTGCGAAGAGTGGATGCTTGGATCC 2156  
Qy 2029 CACATGTTCAGGCCGGGTGTCTGCGTGAGAGGGCAATGCCATTGCTTTGTGGATGGGG 2088  
Db 2157 AACCTGCTCCAGCCATGGTGTCTGTGAACGGAGAGTGTCTATGCAGCCCCGGCTGGGG 2216  
Qy 2089 AGSCACCAACTGCGAGACCCCGAGCCACATGCTTAGACCAGTGTTCAGGCCACGGAAC 2148  
Db 2217 CGGGCTCAACTGCGAGCTGGCGAGGTTCCAGTGGCCAGACCAGTGTAGTGGGCATGGCAC 2276  
Qy 2149 CTTCTCCCGGACACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTT 2208  
Db 2277 TTACCTCCCTGACTCTGGCTCTGCAACTGTGATCCGAATTGGATGGTCCCGACTGCTC 2336  
Qy 2209 TATCGAGATCTGTGCTGCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCG 2268  
Db 2337 TGTGAAGTGTGCTCAGTAGACTGTGGCACTACGGCGTGTGCATCGGGGGAGCCTGCCG 2396  
Qy 2269 CTGCGAGGATGGCTGGATGGGGGAGCCCTGCGACACGGGGCCCTGCCACCCCGCTGTGC 2328  
Db 2397 CTGTGAAGAGGCTGGACAGCGCGGCTTGTGACACAGCGGCTGTGCCACCCCGCTGCAT 2456  
Qy 2329 CGAGCATGGGACCTGCCCGCAGCGCAAGTCGAGTGCAGCCCTGGCTGGAATGGCGAACA 2388  
Db 2457 TGAGACGGGACCTGTAAAGATGGCAAAATGTGAATGCCGAGAGGGCTGGAAATGGTGAACA 2516  
Qy 2389 CTGCACCATCGCTCACTATCTGGATAGGTTAAAGAGGTTGCCCTGGGTTGTGCAA 2448  
Db 2517 CTGCACCATTGAT-----GGTGGCCCTGATTTGTGCAA 2549  
Qy 2449 TGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTG 2508  
Db 2550 CGGTAACGGGAGATGCACATGGGTGAGAACAGTGGCAGTGTGTCTGCGAGACCGGCTG 2609  
Qy 2509 GAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCGCTGCGGTGACAGCAAGACAATGA 2568  
Db 2610 GAGAGGCGCCGGATGCAACCTGTGCCATGGAAACCTCTGCGCTGTATACAAGGATAATGA 2669  
Qy 2569 TGGAGATGGCCTGGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATAT 2628  
Db 2670 GGGAGATGGCCTGGTGGACTGCTGGACCCCTGACTGCTGCCCTCCAGTCACTGCTGCAGAA 2729  
Qy 2629 CAACCCGCTGTGCCCTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCC 2688  
Db 2730 CAGCCTGCTCTGTGCGGGGTCTCGGGACCCCTTGGACATCATTCAGCAAGGCCAGACAGA 2789

Qy 2689 TGTGTACAGCAGAACCTACACTCCTTCTATATACCCGCATCAAGTTCTCCTGTTGGCAGGGA 2748  
Db 2790 CTGGCCTGCGGTGA-----AGTCTTCTATGATCGTATCAAGCTCTTGGCAGGCAAGGA 2843  
Qy 2749 CAGCAGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGGGCACTTGTGTATTCTG 2808  
Db 2844 CAGCACCACATCATCTCTGGAGAACCCCTTCAATAGCAGCCTGGTGTCTCTGTATCCG 2903  
Qy 2809 TGGCCAAAGTATGATGATCAGATGGAACCCCTCTGGTTGGTGTGAACATCAGTTTGTCAA 2868  
Db 2904 AGGCCAAAGTAGTAACACAGGATGGACCCCTCTGGTGGTGTGAATGTGTCTTTTGTCAA 2963  
Qy 2869 TAAACCTCTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAA 2928  
Db 2964 GTACCCAAAATATGGCTACACCATCACTCGCCAGGACGGCACGTTTGAACCTGATTGCCAA 3023  
Qy 2929 TGGCGGCATCTCCATCATCCTGCGGTTTCGAGGGGCACCTTTTCATCACAGGAGCACAC 2988  
Db 3024 TGGGGGCTCTGCCCTTGACTCTTCACTTTGAGCGAGCCCCCTTTCATGAGCCGGAGCGCAC 3083  
Qy 2989 CCTGTGGCTGCCATGGGATCGCTTCTTTGTATGGAACCATCATCATGAGACATGAGGA 3048  
Db 3084 AGTATGGCCGCGGTGGAACAGCTTCTATGCCATGGACACCCCTGGTAAATGAAGACGGAGGA 3143  
Qy 3049 GAATGAGATTCCCAAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTCTCTCTCC 3108  
Db 3144 GAACTCCATCCCAAGCTGTGACCTCAGTGGCTTGTCCGGCCTGATCCGATCATCATCTC 3203  
Qy 3109 ATCCCACTGACGTCCTTCGCCAGCTCCTGTGTCAGAGAAAGGCCCCCATTTGTCGGAAAT 3168  
Db 3204 CTCTCCTCTGTCCACCTTCTCAGCGCTTCCCTCGGGCAACCCCATTTGCTGCTGAGAC 3263  
Qy 3169 TCAGGCTTTGCAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAG 3228  
Db 3264 CCAGGTTCTTCATGAGAGATCGAGCTCCCTGGCACCAACGCTGAAGCTCCGTTACCTCAG 3323  
Qy 3229 CAGCCGGACCCCTGGCTACAAATCTGTCTGAGGATCAGCCTCACCCACCCGACCATCCC 3288  
Db 3324 CTCAGAACACGACAGGTAACAAGTCACTGCTGAAGATCACCATGACCCAGTCCACGGTGCC 3383  
Qy 3289 CTTCAACCTCATGAAGTGCACCTCATGTTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTG 3348  
Db 3384 CTTGAACCTCATCCGGTTCACTTGTATGGTTGCGGTGGAGGGGCATCTCTCCAGAAGTC 3443  
Qy 3349 GTTCGCTGACGCCCAAGACCTGTCTTATTTATTTGGGACAGACAGACAGCTCTACAA 3408  
Db 3444 GTTCCAGGCTTCTCCCAACCTGGCCTACACATTCATCTGGGACAAGACAGACGCTTATGG 3503  
Qy 3409 CCAGAAGGTGTTTGGGCTTTCAGAAAGCTTTGTTTTCGTTGTTATGAATATGAATTCCTG 3468  
Db 3504 CCAAAGGTTTATGGCCTATCGGATGCTGTTGTGTCTGTTGGATTTGAATATGAGACCTG 3563  
Qy 3469 CCCAGATCTAATCCTGTGGGAAAAAAGAAACAACAGTGTGCAGGGCTATGAAATTCACGC 3528  
Db 3564 CCCAGTCTCATCCTGTGGGAAAAAAGGACAGACCCCTACTTCAAGGATTCGAGCTGGACCC 3623  
Qy 3529 GTCCAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTCAACATTCAAAGTGGTAT 3588  
Db 3624 TTCCAACCTTGGTGGCTGGTCCCTGGATAAGCAACACACCCCTCAATGTGAAAAAGCGGAAT 3683  
Qy 3589 CCTGCACAAAGGGAATGGGGAGAACCCAGTTTGTGTCTCAGCAGCCTCCTGTCTATGGGAG 3648  
Db 3684 ACTACTCAAAGGCACAGGGGAGAACCCAGTTCTCGACCCAGCAGCCCGCCCATCATCACCAG 3743  
Qy 3649 CATCATGGCAATGGSCCGGAGAGCATCTCCTGCCCCCAGCTGCAACGGCCTTGTCTGA 3708  
Db 3744 CATCATGGGTAAACGTCGCCCGCAGAGCATCTCCTGTCCAGCTGCAATGGCCTTGTCTGA 3803  
Qy 3709 CGGCAACAAGCTCCTGGCCCCCAGTGGCCCTCACCCTGTGGCTCTGACGGGAGCCTCTATGT 3768  
Db 3804 AGGCAACAACACTGTTGGCCCCCGTGGCCCTGCGCTGTGGGATCGATGGGAGCCTCTTTGT 3863



Db 6021 CAGCATGTCCAGCACACCTCCATTGGCTACATCCGGAACATTTACAACCCACCGGAAAG 6080  
Qy 5986 CAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCACCTCCTTCAACACCTTCTACCT 6045  
Db 6081 CAACGCCCTCGGTCACTCTTTGACTACAGTGATGACGGCCGCATCCTGAAGACGCTTTTCCT 6140  
Qy 6046 GGGCACTGGCCGAGGGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCT 6105  
Db 6141 GGGCACCGGGCGCCAGGTGTTCTATAAGTACGGAAAACTGTCCAAGTTATCGGAGATCGT 6200  
Qy 6106 CTATGACACCAACCAAGGTCAAGTTTCACTATGACGAGACGGCAGGATGCTGAAGACCAT 6165  
Db 6201 CTACGACAGCACTGCCGTCACTTCGGCTATGACGAGACCACTTGGCGTCTGTAAGATGGT 6260  
Qy 6166 CAACCTACAGAATGAGGGCTTCACTGCAACCTGCAACCTACCGTACCGTCAGATTTGGGCCCTGAT 6225  
Db 6261 GAATCTCAAAGCGGGGCTTCTCCTGTACCATCAGGTACCGAAAGGTGGGCCCTCGT 6320  
Qy 6226 TGACCGACAGATCTTCCGCTTCACTGAGGAAGSATGGTCAACGCCCGTTTGTACTACA- 6284  
Db 6321 GGACAAAGCAGATTTACAGTTCTCTGAGGAAGSATGATCAACGCCAGGTTCGATTACAC 6380  
Qy 6285 --ACTATGACAACAGCTTCCGGGTGACCCAGCATGCGAGGCTGTGATCAACGAGACCCCACT 6342  
Db 6381 CTACCACGACAACAGCTTCCGCATCGCCAGCATCAAGCCCGTCACTCAGTGAGACTCCCT 6440  
Qy 6343 GCCCATTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTT 6402  
Db 6441 TCCCGTTGACCTCTACCGGTACGATGAGATTTCTGGCAAGGTGGAACACTTCGGCAAGTT 6500  
Qy 6403 TGGTGTCAATTTACTATGACATTAACAGATCATCAACAGCTGTCTATGACCCACACCAA 6462  
Db 6501 CGGGGTCACTACTACGACATCAACAGATCATCAACCTGCCCCTCATGACACTCAGCAA 6560  
Qy 6463 GCATTTTGTATGATGGAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGTCAAT 6522  
Db 6561 GCATTTTGACACCCATGGCGCATCAAGGAAGTGCAGTATGAGATGTTCCGGTCCCTCAT 6620  
Qy 6523 GTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGGAGCTGAAGGT 6582  
Db 6621 GTACTGGATGACGGTGCAATATGACAGTATGGCAGGGTCACTCAAGAGGGAACCTGAACT 6680  
Qy 6583 AGGACCTTACGCCAATACCACTCGTACTCTCTATGAGTATGATGCTGACGGCCAGCTGCA 6642  
Db 6681 GGGGCCCTATGCCAACACACAAAAGTACACCTATGACTACGACGGGAGCGCCAGCTCCA 6740  
Qy 6643 GACAGTCTCCATCAATGACAAGCCACTCTGGCGTACAGCTACGACCTCAATGGGAACCT 6702  
Db 6741 GAGTGTGGCCGTCAATGACCGGCCCTACCTGGCGTTATAGCTATGACCTCAATGGGAACCT 6800  
Qy 6703 GCACTTACTGAGCCCTGGGAACAGTGCACGGTCAACACCACTACGGTATGACATCCGCGA 6762  
Db 6801 GCACCTGTAAACCCAGGAACAGTGTCTGGCTCATGCCCTCATGCCGTTACGCTATGACCTCCGTGA 6860  
Qy 6763 CCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGCTTCTGAGGCA 6822  
Db 6861 CCGGATAACAGGCTAGGGACGTGCAGTACAAAATCGATGATGATGCTATTTATGCCA 6920  
Qy 6823 GCGGGCGGTGATATCTTTGAGTACAACCTCAGTGGCCCTGCTCATCAAGGCCCTACAACCG 6882  
Db 6921 GAGAGGATCTGACATCTTTGAATACAACCTCCAGGCCCTTCTAACGAGAGCGTACAACAA 6980  
Qy 6883 GGCTGGCAGCTGGAGTGTACGGTACCGCTACGATGGCCCTGGGGCGCGCGTGTCCAGCAA 6942  
Db 6981 GCGCAGCGGTGGAGTGTGCAGTACCGCTATGATGGCGTGAGCCGCGCGGCTTCTTACAA 7040  
Qy 6943 GAGCAGCCACAGCCACCTGCAGTTCTTCTATGACAGACCTGACCAACCCCAACCAAGGT 7002  
Db 7041 GACCAACCTGGGCCACCACTACAGTACTTCTATTCGACCTCCACCAACCCCAACAGTAT 7100  
Qy 7003 CACCCACCTGTACAACCACTCCAGCTCTGAGATCACTTCCCTCTACTACGACTTGCAGG 7062  
Db 7101 CACCCATGTTTACAACCACTCCAACCTCTGAGATCACTCACTCTACTATGACCTCCAGGG 7160

Qy 7063 ACACCTCTTTGCCATGAGAGCTGAGCAGTGGTGATGAGTTTTACATAGCTTGTGACAACAT 7122  
Db 7161 CCACCTCTTTGCCATGAGAGCAGTAGTGGGAAGAGTACTATGTTGCCTCAGATAACAC 7220  
Qy 7123 CGGGACCCCTCTTGTCTCTTTAGTGGAAACAGTTTGTATGATCAAGCAAAATCCTGTACAC 7182  
Db 7221 CGGACTCCTCTGGCTGTTTTTAGTATCAATGGCCTCATGATCAAGCAACTCCAATACAC 7280  
Qy 7183 AGCCTATGGGAGATCTACATGGATACCAACCCCAACTTTTCAGATCATCATAGGCTACCA 7242  
Db 7281 AGCCTATGGGAGATTTACTATGACTCCAATCCAGACTTTTCAGATGGTTCATCGGCTTCCA 7340  
Qy 7243 TGGTGGCCTCTATGATCCACTACCAAGCTTGTCCACATGGGCCGCGAGATTATGATGT 7302  
Db 7341 CGGAGGCCCTCTACGACCCCTCACCAAGCTCGTTCACTTTACGACGGTGTATGACGT 7400  
Qy 7303 GCTGGCCGAGCGTGGACTAGCCAGACCCAGAGCTGTGGAAGCACCTTAGTAGCAGCAA 7362  
Db 7401 GCTGGCAGGACGGTGGACGTCCCCCGACTACACCAATGTGGAGGA---TGTGGGCAAGGA 7457  
Qy 7363 CGTCATGCTTTTAAATCTCTATATGTTCAAAAACAACCCCATCAGCAACTCCCAGGA 7422  
Db 7458 GCCAGCCCCCTTCAACCTGTACATGTTCAAGAAACAACAATCCACTCAGTAATGAGCTGGA 7517  
Qy 7423 CATCAAGTGTCTTATGACAGATGTTAAACAGTGGCTGTCACTTTGGATTCCAGCTACA 7482  
Db 7518 TTTAAAGAACTACGTGACAGACGTGAAGAGCTGGCTCGTGTGTTGGATTTCAGCTCAG 7577  
Qy 7483 CAACGTGATCCCTGGTTATCCAAACCCAGACATGGATGCCATGGAACCCCTCCTACGAGCT 7542  
Db 7578 CAACATCATTCCTGGATTCCCAAGAGCCAAATGTTATTTGTGCTCCGCCCTATGA--- 7634  
Qy 7543 CATCCACACACAGATGAAAAACGAGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACA 7602  
Db 7635 -----ACTGTACAGAGGCCAAGCAAGTGAGAATGGACAGCTCATACAGGTGTCCA 7685  
Qy 7603 GTGTGAAGTACAGAAAGCAGCTCAAGGCCCTTTGTACACCTTAGAACGGTTTGACCACTCTA 7662  
Db 7686 GCAGACAACAGAGAGSCATAACCAAGSCCTTTCTGGCTCTAGAAGGACAGGTCATCTCTAA 7745  
Qy 7663 TGGCTCCACAATCACAGCTGCCAGAGGCTCCAAAGACCAAGAGTTTGCATCCAGCGG 7722  
Db 7745 AAAGTCCATGCAGGCTCCGAGAGAAAGC-----AGGCCACTGGTTGTACGACCAC 7799  
Qy 7723 CTCAGTCTTTGGCAAGGGGTCAAGTTTTCCTTGAAGGATGGCCGAGTGACCAACAGACAT 7782  
Db 7800 GCCCATCATCGGCAAGGCATCATGTTTCGCCATCAAGAAGGGCGGGTGACCAACAGCGT 7859  
Qy 7783 CATCAGTGTGGCCAAATGAGGATGGGCGAAGGGTTGCTGCCATCTTTGAACCAATGCCACTA 7842  
Db 7860 GTCTAGCATCGCCAGTGAAGACAGCCGCAAGGTAGCATCCGTGTTGAACAACGCCCTACTA 7919  
Qy 7843 CCTAGAGAACTGCACTTCAACATTTGATGGGGTGGATACCCATTTACTTTGTGAACCCAGG 7902  
Db 7920 CTTGGACAAGATGCACATACAGCATCGAGGGCAAGGACACACACTACTTCGTGAAGATCGG 7979  
Qy 7903 ACCTTCAGAAGGTGACCTGGCCATCCTGGGCTCAGTGGGGGGCGGCAACCCCTGGAGAA 7962  
Db 7980 TGCAGCGGACGGTGACCTGGTTACGCTGGGACCAACCAATTGGCGCAAGGTGCTGGAGAG 8039  
Qy 7963 TGGGGTCAACGTCATGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTA 8022  
Db 8040 CGGGGTGAACGTCACCGTGTACAGCCCCACGCTGCTGCTGAGCAATACGCTACGGCCTCACCCC 8099  
Qy 8023 CACAGACATCCAGTCCAGTACGGGGCACTGTGCTTGAACACACAGCTACGG----- 8073  
Db 8100 CACCAACATTTGAATTCAGTACTCCAGCTACTCCAGCTGCTGCTCAGCATACGCTACGGCCTCACCCC 8159  
Qy 8074 GACAAAGTTGGATGAGGAGAAGGCACGGGTCTCTGGAGCTGGCCCGGAGAGCCGTGCG 8133  
Db 8160 CGACACACTGGATGAAGAGAAGGCCCGCGCTCTCTGGACCAAGCGCGACAGAGGGCCCTGGG 8219

QY	8134	CCAAGCTGGGCCCCGCGAGCAGCAGAGACTGCGGGAAAGGGAGGAAGGCCTGCGGGCCTG	8193
Db	8220	TACTGCTGGGCCCAAGGAGCAGAGAAAGCCAGGGACCGGAGAGAGGGCAGCCGCTCTGTG	8279
QY	8194	GACAGAGGGGAGAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAAGGCTACGACGGCTT	8253
Db	8280	GACGAGGGCGAGAGCAGCAACTCCTGAGCACGGGACGGGTGCAAGGTTATGAGGGCTA	8339
QY	8254	TTTCGTGATCTCTGTCGAGCAGTACCCAGAACTGTGACACAGCGCCCAACAATCCACTT	8313
Db	8340	TTACGTGCTTCCGGTGGAACAGTACCCAGAGCTGGCAGACAGTAGCAGCAACATCCAGTT	8399
QY	8314	CATGACACAGCGAGATGGGCGGAGGTGACAGA	8348
Db	8400	CTTAGACAGAAATGAGATGGGAAAGAGGTAACAAA	8434

Search completed: June 25, 2004, 03:28:00  
Job time : 31045 secs

**This Page Blank (usp10)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: June 24, 2004, 17:15:34 ; Search time 1916 Seconds  
(without alignments)  
18522.676 Million cell updates/sec

Title: US-10-029-020-13  
Perfect score: 8354  
Sequence: 1 gtttggtgattgtggaggagc.....ccggagggtgacagagaggac 8354  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB	ID	Description			
1	8354	100.0	8354	6	ABS52100	Abs52100	Human	TEN	
2	8225	98.5	8645	6	ABS78652	Abs78652	Human	CDN	
3	7779.2	93.1	8438	6	ABN85378	Abn85378	Human	NOV	
4	3504.6	42.0	8645	6	ABQ82344	Abq82344	Human	NOV	
5	3472.6	41.6	8675	6	ABQ82343	Abq82343	Human	NOV	
6	3424.4	41.0	8473	6	ABQ82345	Abq82345	Human	NOV	
7	3399.8	40.7	8487	6	ABQ82346	Abq82346	Human	NOV	
8	3385	40.5	9826	5	AAS14085	Aas14085	Human	FCT	
9	3362.6	40.3	9826	9	ADB32023	Adb32023	Human	FCT	
10	3351.4	40.1	9729	5	AAS14089	Aas14089	Human	FCT	
11	3351.4	40.1	9729	9	ADB32028	Adb32028	Human	FCT	
12	3274.2	39.2	13202	4	AAK51828	Aak51828	Human	pol	
13	3258	39.0	12879	6	ABK92230	Abk92230	Prostate		
14	3165	37.9	9695	7	ACC72052	Acc72052	BCU0205B		
15	3068.6	36.7	9058	7	ACC72051	Acc72051	BCU0205A		
16	1708.8	20.5	3111	2	AAV19251	Aav19251	Human gam		
17	1708.8	20.5	3111	2	AAx87705	Aax87705	Gamma-her		
18	1708.8	20.5	3111	3	AAD00791	Aad00791	Human Her		
19	1708.8	20.5	3111	4	AAS18526	Aas18526	DNA encod		
20	1562.6	18.7	4245	7	AAL60056	Aal60056	Human PcO		
21	1560.4	18.7	3614	4	AAH14183	Aah14183	Human CDN		
22	1437.4	17.2	3270	4	AAH14671	Aah14671	Human CDN		
23	1193	14.3	3312	4	AAK52812	Aak52812	Human pol		

24	1185.6	14.2	2387	2	AAV19252	Aav19252 Human gam
25	888	10.6	2157	4	AAH14096	Aah14096 Human CDN
26	596.4	7.1	10242	4	ABL29075	AbL29075 Drosophil
27	581.2	7.0	1973	4	AAI14695	Aai14695 Probe #46
28	581.2	7.0	1973	4	ABA56427	Aba56427 Human foe
29	581.2	7.0	1973	4	AAI36063	Aai36063 Probe #47
30	581.2	7.0	1973	4	ABA45900	Aba45900 Human bre
31	581.2	7.0	1973	4	ABA26060	Aba26060 Probe #45
32	581.2	7.0	1973	4	AAK30099	Aak30099 Human bon
33	581.2	7.0	1973	4	AAK04592	Aak04592 Human bra
34	581.2	7.0	1973	4	ABS29749	Abs29749 Human liv
35	581.2	7.0	1973	5	AAI04501	Aai04501 Probe #44
36	581.2	7.0	1973	6	ABS04679	Abs04679 Human gen
37	486.6	5.8	1227	5	AAS68861	Aas68861 DNA encod
38	450.4	5.4	453	3	AAC74626	Aac74626 Human ORF
39	438.4	5.2	2434	5	AAS71215	Aas71215 DNA encod
40	430	5.1	843	4	AAH05415	Aah05415 Human CDN
41	397	4.8	728	5	AAS68860	Aas68860 DNA encod
42	387	4.6	812	4	AAH06979	Aah06979 Human CDN
43	374.6	4.5	640	4	AAAL26459	Aal26459 Human bre
44	373.2	4.5	806	4	AAI23889	Aai23889 Probe #13
45	373.2	4.5	806	4	ABA69008	Aba69008 Human foe

ALIGNMENTS

RESULT 1  
ABS52100  
ID ABS52100 standard; DNA; 8354 BP.  
XX AC ABS52100;  
XX DT 05-NOV-2002 (first entry)  
XX DE Human TEN-M4-like gene.  
XX KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;  
KW cell signal processing; metabolic pathway modulation; metabolic disorder;  
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;  
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;  
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;  
KW memory defect; infertility; congenital heart defect; hair growth;  
KW pigmentation disorder; endocrine disorder; respiratory disease; health;  
KW gastro-intestinal disease; reproductive; neurological disease;  
KW bone marrow transplantation; endocrine disease; allergy; inflammation;  
KW nephrological disorder; urinary system disorder; age-related disorder;  
KW neuropsychiatric disorder; EGF-related protein; SCUBEL; TEN-M4;  
KW adipocyte complement-related Clq tumour necrosis factor; out at first;  
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;  
KW type Ia membrane sushi-containing domain; butyrophilin;  
KW type Ia membrane-sushi domain containing; SNP; gene; ds;  
KW single nucleotide polymorphism.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
XX variation replace(117,G)  
FT /\*tag= a  
FT /standard\_name= "Single nucleotide polymorphism (SNP)"  
FT replace(225,C)  
FT /\*tag= b  
FT /standard\_name= "Single nucleotide polymorphism (SNP)"  
FT replace(260,A)  
FT /\*tag= c  
FT /standard\_name= "Single nucleotide polymorphism (SNP)"  
XX WO200257453-A2.  
XX PN 25-JUL-2002.  
XX PD 19-DEC-2001; 2001WO-US050331.  
XX PF XX



QY 1441 TCTGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCTCCCTCCTTACATAC 1500  
Db 1441 TCTGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCTCCCTCCTTACATAC 1500  
QY 1501 ACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGSAGGCTCCTAACCCAGAGGCGCGGAG 1560  
Db 1501 ACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGSAGGCTCCTAACCCAGAGGCGCGGAG 1560  
QY 1561 CCTAGAGGGGACCCCGCGCCAGTCTCGGGGAACTGTGCCCCCCTCCAGCCATGAGACAGG 1620  
Db 1561 CCTAGAGGGGACCCCGCGCCAGTCTCGGGGAACTGTGCCCCCCTCCAGCCATGAGACAGG 1620  
QY 1621 CTTTATCCAGTATTTGGATTTCAGGAATCTGGCAGCTGGCTTTTACAAATGACGGAAGGA 1680  
Db 1621 CTTTATCCAGTATTTGGATTTCAGGAATCTGGCAGCTGGCTTTTACAAATGACGGAAGGA 1680  
QY 1681 GTCAGAAAGTGGTTTCCTTTCTCACCACCTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA 1740  
Db 1681 GTCAGAAAGTGGTTTCCTTTCTCACCACCTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA 1740  
QY 1741 CTGCTATGGCAATGGTGATGCGATCTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGG 1800  
Db 1741 CTGCTATGGCAATGGTGATGCGATCTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGG 1800  
QY 1801 CCCCAGCTGTGGCAGAGCCTCCTGCTGCCCTGCTGTAGCGGAAATGGCCAAATACATGAA 1860  
Db 1801 CCCCAGCTGTGGCAGAGCCTCCTGCTGCCCTGCTGTAGCGGAAATGGCCAAATACATGAA 1860  
QY 1861 AGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGTGAAGTCGATGTGCCACCAACCA 1920  
Db 1861 AGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGTGAAGTCGATGTGCCACCAACCA 1920  
QY 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCAGGGCACCTGCATCTGCAA 1980  
Db 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCAGGGCACCTGCATCTGCAA 1980  
QY 1981 CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGAACCCACATGTTTCA 2040  
Db 1981 CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGAACCCACATGTTTCA 2040  
QY 2041 CCGGGGTGCTCGCTGAGAGGCGAATGCCATTGCTTTTGTGGATGGGAGGCGCAACTG 2100  
Db 2041 CCGGGGTGCTCGCTGAGAGGCGAATGCCATTGCTTTTGTGGATGGGAGGCGCAACTG 2100  
QY 2101 CGAGACCCCGAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAACTTCTCCCGGA 2160  
Db 2101 CGAGACCCCGAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAACTTCTCCCGGA 2160  
QY 2161 CACCGGGCTTTGCAGCTGTGACCCAAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG 2220  
Db 2161 CACCGGGCTTTGCAGCTGTGACCCAAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG 2220  
QY 2221 TGCTGCCGACTGTGGTGCCATGGCGTGTGCGTAGGGGGACCTGCCGCTGGAGGATGG 2280  
Db 2221 TGCTGCCGACTGTGGTGCCATGGCGTGTGCGTAGGGGGACCTGCCGCTGGAGGATGG 2280  
QY 2281 CTGGATGGGGCAGCCTGCGACACAGCGGGCCTGCCACCCCGCTGTGCCGAGCATGGGAC 2340  
Db 2281 CTGGATGGGGCAGCCTGCGACACAGCGGGCCTGCCACCCCGCTGTGCCGAGCATGGGAC 2340  
QY 2341 CTGCCGCGACGGCAAGTCGAGTGCAGCCTGGCTGGAAATGGCGAACTGCAACATCGC 2400  
Db 2341 CTGCCGCGACGGCAAGTCGAGTGCAGCCTGGCTGGAAATGGCGAACTGCAACATCGC 2400  
QY 2401 TCACATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAG 2460  
Db 2401 TCACATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAG 2460  
QY 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGG 2520  
Db 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGG 2520  
QY 2521 CTGTGACACTTCCATGGAGACTGCTCGGGTGACAGCAAAAGACAATGATGGAGATGSCCT 2580

Db 2521 CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGSCCT 2580  
QY 2581 GGTGACTGATGGACCCCTGACTGCTGCCTCCAGCCCCCTGTGCCATATCAACCCCGCTGTG 2640  
Db 2581 GGTGACTGATGGACCCCTGACTGCTGCCTCCAGCCCCCTGTGCCATATCAACCCCGCTGTG 2640  
QY 2641 CTTTGGCTCCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTCAAGCA 2700  
Db 2641 CTTTGGCTCCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTCAAGCA 2700  
QY 2701 GAACCTACACTCCTTCTATGACCGCATCAAGTTCTCTGTGGCAGGGACAGCACGCACAT 2760  
Db 2701 GAACCTACACTCCTTCTATGACCGCATCAAGTTCTCTGTGGCAGGGACAGCACGCACAT 2760  
QY 2761 AATCCCGGGGAGAAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGCCAAAGTGAT 2820  
Db 2761 AATCCCGGGGAGAAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGCCAAAGTGAT 2820  
QY 2821 GACATCAGATGGAACCCCTCGTGTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTT 2880  
Db 2821 GACATCAGATGGAACCCCTCGTGTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTT 2880  
QY 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGCGGCACTCTC 2940  
Db 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGCGGCACTCTC 2940  
QY 2941 CATCATCCTCGGTTTCGAGCGGCGACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCC 3000  
Db 2941 CATCATCCTCGGTTTCGAGCGGCGACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCC 3000  
QY 3001 ATGGGATCGCTTCTTGTCTATGGAACCATCATCATGAGACATGAGGAGATGAGATTCC 3060  
Db 3001 ATGGGATCGCTTCTTGTCTATGGAACCATCATCATGAGACATGAGGAGATGAGATTCC 3060  
QY 3061 CAGCTGTGACCTGAGCAATTTTGGCCGCCCCCAACCCAGTCGTCCTCCATCCCCACTGAC 3120  
Db 3061 CAGCTGTGACCTGAGCAATTTTGGCCGCCCCCAACCCAGTCGTCCTCCATCCCCACTGAC 3120  
QY 3121 GTCCTTCCCGCAGCTCCTGTGAGAGAAAGGCCCCATTGTGCGGAAATTCAGGCTTTTGCA 3180  
Db 3121 GTCCTTCCCGCAGCTCCTGTGAGAGAAAGGCCCCATTGTGCGGAAATTCAGGCTTTTGCA 3180  
QY 3181 GGAGGAAATCTCTATCTCTGGTGTCAAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCC 3240  
Db 3181 GGAGGAAATCTCTATCTCTGGTGTCAAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCC 3240  
QY 3241 TGGCTACAAATCTGTCTCTGAGGATCAGCCCTCACCCACCCCGACCATCCCCCTTCAACCTCAT 3300  
Db 3241 TGGCTACAAATCTGTCTCTGAGGATCAGCCCTCACCCACCCCGACCATCCCCCTTCAACCTCAT 3300  
QY 3301 GAAGGTGCACCTCATGGTAGCGGTGAGGGCGCGCTTTCAGGAAAGTGGTTCGCTGCAGC 3360  
Db 3301 GAAGGTGCACCTCATGGTAGCGGTGAGGGCGCGCTTTCAGGAAAGTGGTTCGCTGCAGC 3360  
QY 3361 CCCAGACCTGCTCTATTATTTCATTTGGGACAAAGACAGACGCTGTACAAACAGAGGTGTT 3420  
Db 3361 CCCAGACCTGCTCTATTATTTCATTTGGGACAAAGACAGACGCTGTACAAACAGAGGTGTT 3420  
QY 3421 TGGGCTTTCAGAAAGCCTTGTGTTCCGTGGTTATGAATATGAATCCTGCCAGATCTAAT 3480  
Db 3421 TGGGCTTTCAGAAAGCCTTGTGTTCCGTGGTTATGAATATGAATCCTGCCAGATCTAAT 3480  
QY 3481 CCTGTGGGAAAAAGAAACAACAGTGTGCAGGGCTATGAAATTGACCGCTCCAAGCTTGG 3540  
Db 3481 CCTGTGGGAAAAAGAAACAACAGTGTGCAGGGCTATGAAATTGACCGCTCCAAGCTTGG 3540  
QY 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTGCACAAAG 3600  
Db 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTGCACAAAG 3600  
QY 3601 GAATGGGAGAACCAAGTTTGTGTTCTCAGCAGCCCTCCTGTCTATTTGGAGCATCATGGGCAA 3660

Db 3601 GAATGGGGAGAACAGTTTGTGTCTCAGCAGCCCTCTGTCTATTGGGAGCATCATGGGGCAA 3660  
 Qy 3661 TGGGGCCCGGAGAAGCATCTCCTGCCCCAGCTGGAACGGCCCTTGCTGACGGCAACAAGCT 3720  
 Db 3661 TGGGGCCCGGAGAAGCATCTCCTGCCCCAGCTGGAACGGCCCTTGCTGACGGCAACAAGCT 3720  
 Qy 3721 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCCTCTATGTGGGTGATTTCAA 3780  
 Db 3721 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCCTCTATGTGGGTGATTTCAA 3780  
 Qy 3781 CTACATTAGAAGGATCTTCCCCTCTGGAATGTCAACCAATCTTAGAGTGAGGAATAA 3840  
 Db 3781 CTACATTAGAAGGATCTTCCCCTCTGGAATGTCAACCAATCTTAGAGTGAGGAATAA 3840  
 Qy 3841 AGATTTAGACATAGTCACAGTCCAGCACACAATACTACCTGGCCACAGACCCCATGAG 3900  
 Db 3841 AGATTTAGACATAGTCACAGTCCAGCACACAATACTACCTGGCCACAGACCCCATGAG 3900  
 Qy 3901 TGGGGCCGTCTTCCCTTCTGACAGCAACAGCCGGGGTCTTTAAATCAAGTCCACTGT 3960  
 Db 3901 TGGGGCCGTCTTCCCTTCTGACAGCAACAGCCGGGGTCTTTAAATCAAGTCCACTGT 3960  
 Qy 3961 GGTGGTGAAGGACCTTTGCAAGAACTCTGAGGTGGTTCGGGGACAGGTGACCATGTCCT 4020  
 Db 3961 GGTGGTGAAGGACCTTTGCAAGAACTCTGAGGTGGTTCGGGGACAGGTGACCATGTCCT 4020  
 Qy 4021 CCCCTTTGATGACACTCGCTGCGGGATGTTGGAAGGCCACAGAGCCACACTCACCAA 4080  
 Db 4021 CCCCTTTGATGACACTCGCTGCGGGATGTTGGAAGGCCACAGAGCCACACTCACCAA 4080  
 Qy 4081 TCCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT 4140  
 Db 4081 TCCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT 4140  
 Qy 4141 CAGACGCATCGATCAGAAATGGGATCATCTCACCCCTGCTCGGCTCTAATGATCTCACATC 4200  
 Db 4141 CAGACGCATCGATCAGAAATGGGATCATCTCACCCCTGCTCGGCTCTAATGATCTCACATC 4200  
 Qy 4201 AGCCCGGCCACTCAGCTGTGATTTCTGTCAATGGATATTTCCAGGTAAAGACTGGAGTGGCC 4260  
 Db 4201 AGCCCGGCCACTCAGCTGTGATTTCTGTCAATGGATATTTCCAGGTAAAGACTGGAGTGGCC 4260  
 Qy 4261 CACAGACTTAGCCATCAACCCCAATGGACAACTCACTTTATGTCTCTCGACAACTATGTGGT 4320  
 Db 4261 CACAGACTTAGCCATCAACCCCAATGGACAACTCACTTTATGTCTCTCGACAACTATGTGGT 4320  
 Qy 4321 CCTGCAAAATCTCTGAAAACCAACAGGTGCGCATTTGTCGGGGAGGCCCATGCACTGCCA 4380  
 Db 4321 CCTGCAAAATCTCTGAAAACCAACAGGTGCGCATTTGTCGGGGAGGCCCATGCACTGCCA 4380  
 Qy 4381 GGTCCCTGGCATTGACCACTTCTGTCAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440  
 Db 4381 GGTCCCTGGCATTGACCACTTCTGTCAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440  
 Qy 4441 AGCCACCGCTTTGGCTGTTTACACAATGGGTCCTGTATATTGTGAGACTGATGAGAA 4500  
 Db 4441 AGCCACCGCTTTGGCTGTTTACACAATGGGTCCTGTATATTGTGAGACTGATGAGAA 4500  
 Qy 4501 AAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTGGGGC 4560  
 Db 4501 AAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTGGGGC 4560  
 Qy 4561 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTGTTTTCTGGAGACGATGG 4620  
 Db 4561 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTGTTTTCTGGAGACGATGG 4620  
 Qy 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGGA 4680  
 Db 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGGA 4680  
 Qy 4681 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCCTTT 4740  
 Db 4681 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCCTTT 4740

Qy 4741 CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAGGAGCTCTATCTGTT 4800  
 Db 4741 CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAGGAGCTCTATCTGTT 4800  
 Qy 4801 TGATACCAACCGGCAAGCACCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAA 4860  
 Db 4801 TGATACCAACCGGCAAGCACCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAA 4860  
 Qy 4861 CTTTACCTTACACTGGGACGGGACATCATCACTCATCAAGACAACATGGCAACATGGT 4920  
 Db 4861 CTTTACCTTACACTGGGACGGGACATCATCACTCATCAAGACAACATGGCAACATGGT 4920  
 Qy 4921 AAATGTCCCGGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGT 4980  
 Db 4921 AAATGTCCCGGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGT 4980  
 Qy 4981 GTACTGGGTGACCATGGGACCAACAGTGTCACTCAAGAGTGTGACACACAAGGACACGA 5040  
 Db 4981 GTACTGGGTGACCATGGGACCAACAGTGTCACTCAAGAGTGTGACACACAAGGACACGA 5040  
 Qy 5041 GTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAA 5100  
 Db 5041 GTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAA 5100  
 Qy 5101 CGGATGGACAACATTTATGAGTACGACAGTTTGGCCCTTGGCAACCAAAAGCAATGAAA 5160  
 Db 5101 CGGATGGACAACATTTATGAGTACGACAGTTTGGCCCTTGGCAACCAAAAGCAATGAAA 5160  
 Qy 5161 TACTGGCCAGGTGAGCAGTTTCCGAAAGTGATACAGACAGTTTCAAGTGTCCAGGTAGA 5220  
 Db 5161 TACTGGCCAGGTGAGCAGTTTCCGAAAGTGATACAGACAGTTTCAAGTGTCCAGGTAGA 5220  
 Qy 5221 GACCTCCAGCAAGGATGATGTACCAATAACCAACCACTGTCTGCTCAGGCGCTTCTA 5280  
 Db 5221 GACCTCCAGCAAGGATGATGTACCAATAACCAACCACTGTCTGCTCAGGCGCTTCTA 5280  
 Qy 5281 CACACTGCTGCAAGCAAGTCCGGAACAGTACTACATCGGGCCGATGGCTCCTTGGC 5340  
 Db 5281 CACACTGCTGCAAGCAAGTCCGGAACAGTACTACATCGGGCCGATGGCTCCTTGGC 5340  
 Qy 5341 GCTGCTGTGGCCCAACGGCATGGAGGTGGCGTGCAGACTGAGCCCACTTGTGGCTGG 5400  
 Db 5341 GCTGCTGTGGCCCAACGGCATGGAGGTGGCGTGCAGACTGAGCCCACTTGTGGCTGG 5400  
 Qy 5401 CACCGTCAACCCCGACCGTGGGCAAGAGGAAATGTACGCTGCCCATCGACAACGGCCTCAA 5460  
 Db 5401 CACCGTCAACCCCGACCGTGGGCAAGAGGAAATGTACGCTGCCCATCGACAACGGCCTCAA 5460  
 Qy 5461 CCTGGTGGAGTGGCCCAACGGCAAGAGAGGCTCGGGGCCAGGTCACTGTCTTTGGGCG 5520  
 Db 5461 CCTGGTGGAGTGGCCCAACGGCAAGAGAGGCTCGGGGCCAGGTCACTGTCTTTGGGCG 5520  
 Qy 5521 CCGGCTGGGGTGCACAACCGAAATCTCTATCTCTGGAATTTGATCGCGTAACACGCAC 5580  
 Db 5521 CCGGCTGGGGTGCACAACCGAAATCTCTATCTCTGGAATTTGATCGCGTAACACGCAC 5580  
 Qy 5581 AGAGAAGATCTATGATGACCAACCGCAAGTCACTTCCGATTTCTGATCGACAGGCGGG 5640  
 Db 5581 AGAGAAGATCTATGATGACCAACCGCAAGTCACTTCCGATTTCTGATCGACAGGCGGG 5640  
 Qy 5641 GCGGCCAGCCTCTGGTCAACCCAGCAGGAGCTGAATGGTGTCAACGCTGACATACTCCCC 5700  
 Db 5641 GCGGCCAGCCTCTGGTCAACCCAGCAGGAGCTGAATGGTGTCAACGCTGACATACTCCCC 5700  
 Qy 5701 TGGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCA 5760  
 Db 5701 TGGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCA 5760  
 Qy 5761 GCGGGCCGCATCACATCCAGGATCTTCCGCTGATGGGAAGACATGGAGCTACACATCTT 5820  
 Db 5761 GCGGGCCGCATCACATCCAGGATCTTCCGCTGATGGGAAGACATGGAGCTACACATCTT 5820



Db 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8040  
Qy 8041 GTACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTGGATGAGAGAAGGCACG 8100  
Db 8041 GTACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTGGATGAGAGAAGGCACG 8100  
Qy 8101 GGTCTGGAGCTGGCCCGCAGAGAGCCGTGCGCAAGCGTGGGCCCGCAGCAGCAGAG 8160  
Db 8101 GGTCTGGAGCTGGCCCGCAGAGAGCCGTGCGCAAGCGTGGGCCCGCAGCAGCAGAG 8160  
Qy 8161 ACTCGGGGAAGGGAGGAGGCGCTGCGGCGCTGGACAGAGGGGGAGAGCAGGCTGCT 8220  
Db 8161 ACTCGGGGAAGGGAGGAGGCGCTGCGGCGCTGGACAGAGGGGGAGAGCAGGCTGCT 8220  
Qy 8221 GAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCC 8280  
Db 8221 GAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCC 8280  
Qy 8281 AGAACTGTCAGACAGCGGCCCAACAACATCCATTCATGAGACAGAGCGGAGATGGCCGGAG 8340  
Db 8281 AGAACTGTCAGACAGCGGCCCAACAACATCCATTCATGAGACAGAGCGGAGATGGCCGGAG 8340  
Qy 8341 GTGACAGAGAGGAC 8354  
Db 8341 GTGACAGAGAGGAC 8354

RESULT 2  
ABS78652

ID ABS78652 standard; cDNA; 8645 BP.

AC ABS78652;

XX 16-DEC-2002 (first entry)

XX Human cDNA encoding CGDD10, INCYTE 7488573CB1.

KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;  
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;  
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;  
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;  
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;  
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;  
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;  
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;  
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;  
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;  
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;  
KW rheumatoid arthritis.

XX Homo sapiens.

OS WO200272830-A2.

XX 19-SEP-2002.

PF 08-FEB-2002; 2002WO-US003715.

XX 09-FEB-2001; 2001US-0268111P.

PR 23-FEB-2001; 2001US-0271175P.

PR 08-MAR-2001; 2001US-0274503P.

PR 09-MAR-2001; 2001US-0274552P.

XX (INCY-) INCYTE GENOMICS INC.

PA Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;

PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;

PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;

XX WPI; 2002-723356/78.

DR P-PSDB; ABG97359.

XX New human proteins associated with cell growth, differentiation and

PT death, useful for diagnosing, treating or preventing autoimmune or  
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,  
PT atherosclerosis or hepatitis.

XX Claim 5; Page 175-178; 181pp; English.

XX The invention relates to an isolated polypeptide comprising CGDD1-12  
CC (cell growth, differentiation and death), a naturally occurring amino  
CC acid sequence at least 90% identical to CGDD, a biologically active  
CC fragment or an immunogenic fragment. Also included are the  
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide  
CC comprising a promoter sequence operably linked to the CGDD  
CC polynucleotides, a cell transgene formed with the recombinant polynucleotide,  
CC a transgenic organism comprising the recombinant polynucleotide, an anti-  
CC CGDD antibody, screening for compounds which bind to/modulate or are  
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a  
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,  
CC agonists and antagonists are useful for diagnosing, treating or  
CC preventing disorders associated with aberrant expression of CGDD,  
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,  
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia  
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental  
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),  
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or  
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in  
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,  
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune  
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,  
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,  
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,  
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic  
CC infections. They are also useful in the assessment of the effects of  
CC exogenous compounds on the expression of nucleic acid and amino acid  
CC sequences of proteins associated with CGDD. The present sequence encodes  
CC a CGDD protein

XX Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

SQ Query Match 98.5%; Score 8225; DB 6; Length 8645;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 8300; Conservative 0; Mismatches 21; Indels 33; Gaps 3;

Qy 1 GTTTGTGATGTGGAGAGCGCGGGCGGAGCCCATGGACGTGAAGAGAGAGCCTTA 60

Db 83 GTTTGTGATGTGGAGAGCGCGGGCGGAGCCCATGGACGTGAAGAGAGAGCCTTA 142

Qy 61 CCGCTCGTGTACCCGCGCGCGAGCGCGGAGCGCGCTACACCGCTCGTCCGCGGACAG 120

Db 143 CCGCTCGTGTACCCGCGCGCGAGCGCGGAGCGCGCTACACCGCTCGTCCGCGGACAG 202

Qy 121 CGAGGAGGGCAAGCCCCCGCAGAAATCGTACAGCTCCAGCGAGACCTGAAGGCCTACGA 180

Db 203 CGAGGAGGGCAAGCCCCCGCAGAAATCGTACAGCTCCAGCGAGACCTGAAGGCCTACGA 262

Qy 181 CCAGGACGCCCCCGCTAGCCTATGGCAGCCCGCTCAAGGACATTGTCCGCGAGGCGCGA 240

Db 263 CCAGGACGCCCCCGCTAGCCTATGGCAGCCCGCTCAAGGACATTGTCCGCGAGGCGCGA 322

Qy 241 GGAATTCTGCCGCACAGGTGCCAACTTCAACCCTGCGGGAGCTGGGGCTGGAAGAATAAC 300

Db 323 GGAATTCTGCCGCACAGGTGCCAACTTCAACCCTGCGGGAGCTGGGGCTGGAAGAATAAC 382

Qy 301 GCCCCCTCACGGGACCCCTGTACCGGACAGACATTGGCCCTGCCCCAATGCGGGCTACTCCAT 360

Db 383 GCCCCCTCACGGGACCCCTGTACCGGACAGACATTGGCCCTGCCCCAATGCGGGCTACTCCAT 442

Qy 361 GGGGGCTGGCTCTGATGCCGACATGGAGCTGACACGGTGTCTCCCTGAGCACCCCGT 420

Db 443 GGGGGCTGGCTCTGATGCCGACATGGAGCTGACACGGTGTCTCCCTGAGCACCCCGT 502

Qy 421 GCGTCTGTGGGCGCGGAGCACACGGTACGGGCGCAGCTCCTGCTGTCCAGCCGGCCAA 480

Db 503 GCGTCTGTGGGCGCGGAGCACACGGTACGGGCGCAGCTCCTGCTGTCCAGCCGGCCAA 562

QY	481	TTCCAATCTCACACTCACCGACACCGAGCATGAAAACACTGAGACTGATCATCCGGCGG	540
Db	563	TTCCAATCTCACACTCACCGACACCGAGCATGAAAACACTGAGAC-----TCCGGGCGG	616
QY	541	CTGCAGAAACCAACGCGGGCTCCGGACGCCGCCGCCGCGCTCTCGCACGCCACACACCC	600
Db	617	CTGCAGAAACCAACGCGGGCTCCGGACGCCGCCGCCGCGCTCTCGCACGCCACACACCC	676
QY	601	CAACCAGCACCAACGCGGCTCCATTAACTCCCTGAACCGGGGCAACTTCACGCCGAGGAG	660
Db	677	CAACCAGCACCAACGCGGCTCCATTAACTCCCTGAACCGGGGCAACTTCACGCCGAGGAG	736
QY	661	CAACCCAGCCCGCCCCCACGGACCACTCGCTCTCCGAGAGCCCCCTGCCGGCGGCGC	720
Db	737	CAACCCAGCCCGCCCCCACGGACCACTCGCTCTCCGAGAGCCCCCTGCCGGCGGCGC	796
QY	721	CCAGGAGCCTGCCACCGCCAGGAGAACTGGCTGCTCAACAGCAACATCCCCCTGGAGAC	780
Db	797	CCAGGAGCCTGCCACCGCCAGGAGAACTGGCTGCTCAACAGCAACATCCCCCTGGAGAC	856
QY	781	CAGGAACTTAGGCAAGCAGCCATTCTTAGGACATTGCAGGACAACTCATTGAGATGGA	840
Db	857	CAGAAACTTAGGCAAGCAGCCATTCTTAGGACATTGCAGGACAACTCATTGAGATGGA	916
QY	841	CATTCTGGCGCCTCCGCCATGATGGGCTTACAGTACGGGCACTTCTCTTCAAGCC	900
Db	917	CATTCTGGCGCCTCCGCCATGATGGGCTTACAGTACGGGCACTTCTCTTCAAGCC	976
QY	901	TGGAGGCACCTCCCCGCTCTTCTGCACCAATCACAGGGTACCCACTGACGTCCAGCAC	960
Db	977	TGGAGGCACCTCCCCGCTCTTCTGCACCAATCACAGGGTACCCACTGACGTCCAGCAC	1036
QY	961	AGTGTACTCTCTCCGCCCGACCCCTGCCCGCAGACACTTCGCCCGCGCGGCTTTAA	1020
Db	1037	AGTGTACTCTCTCCGCCCGACCCCTGCCCGCAGACACTTCGCCCGCGCGGCTTTAA	1096
QY	1021	CCTCAAGAAAGCCCTCCAAGTACTGTAAGTGAAGTGGCAGCCCTGAGCGCATCGTCAT	1080
Db	1097	CCTCAAGAAAGCCCTCCAAGTACTGTAAGTGAAGTGGCAGCCCTGAGCGCATCGTCAT	1156
QY	1081	CTCAGCCACTCTGGTCACTCCTGCTGGCATACTTTGTGGCCATGCACCTGTTGGCCTAAA	1140
Db	1157	CTCAGCCACTCTGGTCACTCCTGCTGGCATACTTTGTGGTAAAGCACT-----CTTCAA	1210
QY	1141	CTGGCACCTGCAGCCGATGAGGGGCAGATGTATGAGATCACGGAGGACACAGCCAGCAG	1200
Db	1211	CTGGCACCTGCAGCCGATGAGGGGCAGATGTATGAGATCACGGAGGACACAGCCAGCAG	1270
QY	1201	TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACCTGGCTTAGAGACCCC	1260
Db	1271	TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACCTGGCTTAGAGACCCC	1330
QY	1261	TGACAGGAAAGGCAAGGAACCAAGAGGAAAGCCAGTAGTTTCTTTTCCAGAGGACAG	1320
Db	1331	TGACAGGAAAGGCAAGGAACCAAGAGGAAAGCCAGTAGTTTCTTTTCCAGAGGACAG	1390
QY	1321	TTTTCATAGATTCTGGAGAAATTGATGTGGGAAGGCGAGCCTCCCAGAAGATTCTCTCTGG	1380
Db	1391	TTTTCATAGATTCTGGAGAAATTGATGTGGGAAGGCGAGCTTCCCAGAAGATTCTCTCTGG	1450
QY	1381	CACTTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCATCTGAAATTCAATGTGTC	1440
Db	1451	CACTTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCATCTGAAATTCAATGTGTC	1510
QY	1441	TCTGGGAAAGGCAGCCCTGGTTGGCATTTTATGGCAGAAAAGGCCCTCCCTCTTACATAC	1500
Db	1511	TCTGGGAAAGGCAGCCCTGGTTGGCATTTTATGGCAGAAAAGGCCCTCCCTCTTACATAC	1570
QY	1501	ACAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAG	1560
Db	1571	ACAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAG	1630

Qy	1561	CCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTGCCCCCCTCCAGCCATGAGACAGG	1620
Db	1631	CCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTGCCCCCCTCCAGCCATGAGACAGG	1690
Qy	1621	CTTCATCCAGTATTTGGATTCAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGA	1680
Db	1691	CTTCATCCAGTATTTGGATTCAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGA	1750
Qy	1681	GTCAGAAAGTGGTTTCCTTTCTCACCACTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA	1740
Db	1751	GTCAGAAAGTGGTTTCCTTTCTCACCACTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA	1810
Qy	1741	CTGCTATGGCAATGGTGACTGCATCTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGG	1800
Db	1811	CTGCTATGGCAATGGTGACTGCATCTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGG	1870
Qy	1801	CCCCGACTGTGGCAGAGCCTCTCTGCCCCGCTGCTCTGTAGCGGAAATGGCCCAATACATGAA	1860
Db	1871	CCCCGACTGTGGCAGAGCCTCTCTGCCCCGCTGCTCTGTAGCGGAAATGGCCCAATACATGAA	1930
Qy	1861	AGGCAGATGCTTGIGCCACAGTGGCTGGAAAGSGCGCTGAGTGCATGTGCCCAACCAACCA	1920
Db	1931	AGGCAGATGCTTGIGCCACAGTGGCTGGAAAGSGCGCTGAGTGCATGTGCCCAACCAACCA	1990
Qy	1921	GTGTATCGATGTGGCCTGCAGCAACCATGGCACTGCATCACGGGCACCTGCATCTGCAA	1980
Db	1991	GTGTATCGATGTGGCCTGCAGCAACCATGGCACTGCATCACGGGCACCTGCATCTGCAA	2050
Qy	1981	CCCTGGCTACAAGGCGAGAGCTGTAGGAAGTGGACTGCATGGACCCCAATGTTTCAGG	2040
Db	2051	CCCTGGCTACAAGGCGAGAGCTGTAGGAAGTGGACTGCATGGACCCCAATGTTTCAGG	2110
Qy	2041	CCGGGGTGTCTGCTGAGAGGCGAATGCCATTGCTTTGTGGGATGGGAGGCACCAACTG	2100
Db	2111	CCGGGGTGTCTGCTGAGAGGCGAATGCCATTGCTTTGTGGGATGGGAGGCACCAACTG	2170
Qy	2101	CGAGACCCCAAGGCCACATGCTTAGACCAAGTGTTCAGGCCACCGAACCTTCTCCCGGA	2160
Db	2171	CGAGACCCCAAGGCCACATGCTTAGACCAAGTGTTCAGGCCACCGAACCTTCTCCCGGA	2230
Qy	2161	CACCGGCTTTGCAGTGTGACCCCAAGCTGGAATGGACACGACTGTTCTATCGAGATCTG	2220
Db	2231	CACCGGCTTTGCAGTGTGACCCCAAGCTGGAATGGACACGACTGTTCTATCGAGATCTG	2290
Qy	2221	TGCTGCCGACTGTGTGGCCCATGGCGTGTGCTGAGGGGGCACCTGCCGCTGGAGGATGG	2280
Db	2291	TGCTGCCGACTGTGTGGCCCATGGCGTGTGCTGAGGGGGCACCTGCCGCTGGAGGATGG	2350
Qy	2281	CTGGATGGGGCAGCTGCGACCAAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGAC	2340
Db	2351	CTGGATGGGGCAGCTGCGACCAAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGAC	2410
Qy	2341	CTGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGAAATGGCGAACACTGCACCATCGC	2400
Db	2411	CTGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGAAATGGCGAACACTGCACCATCGC	2470
Qy	2401	TCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTTGGTTGTGCAATGGCAACGGCAG	2460
Db	2471	TCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTTGGTTGTGCAATGGCAACGGCAG	2530
Qy	2461	ATGTACCTTAGACCTGAATGGTTGGCACTGGCTGCCAGCTGGGCTGGAGAGGAGCTGG	2520
Db	2531	ATGTACCTTAGACCTGAATGGTTGGCACTGGCTGCCAGCTGGGCTGGAGAGGAGCTGG	2590
Qy	2521	CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCT	2580
Db	2591	CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCT	2650
Qy	2581	GGTGGACTGCATGGACCTGACTGCTGCTCCAGCCCCCTGTGCCATATCAACCCCGTGTG	2640
Db	2651	GGTGGACTGCATGGACCTGACTGCTGCTCCAGCCCCCTGTGCCATATCAACCCCGTGTG	2710
Qy	2641	CCTTGGCTCCCCTAAACCTCTGGACATATCCAGGAGACACAGGTCCCTGTGTACAGCA	2700

Db 2711 CCTTGGCTCCCTAACCTCTGGACATCATCCAGGAGACACAGGTCCTCTGTGTACAGCA 2770  
Qy 2701 GAACTTACACTCCTTCTATGACCGCATCAAGTTCTCTGTGGCAGGAGACGACACGACAT 2760  
Db 2771 GAACTTACACTCCTTCTATGACCGCATCAAGTTCTCTGTGGCAGGAGACGACACGACAT 2830  
Qy 2761 AATCCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCTGTGGCCAAAGTGAT 2820  
Db 2831 AATCCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCTGTGGCCAAAGTGAT 2890  
Qy 2821 GACATCAGATGGAAACCCCTCTGTTGGTGTGAACATCAGTTTGTCAATAAACCCCTCTCTT 2880  
Db 2891 GACATCAGATGGAAACCCCTCTGTTGGTGTGAACATCAGTTTGTCAATAAACCCCTCTCTT 2950  
Qy 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC 2940  
Db 2951 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC 3010  
Qy 2941 CATCATCTCGGTTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGTGCC 3000  
Db 3011 CATCATCTCGGTTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGTGCC 3070  
Qy 3001 ATGGGATCGCTTCTTGTGTCATGGAAACCATCATCATGAGACATGAGGAGAATGAGATTCC 3060  
Db 3071 ATGGGATCGCTTCTTGTGTCATGGAAACCATCATCATGAGACATGAGGAGAATGAGATTCC 3130  
Qy 3061 CAGTGTGACCTGAGCAATTTGCCCGCCCCAACCCAGTCGTCCTCTCCATCCCCACTGAC 3120  
Db 3131 CAGTGTGACCTGAGCAATTTGCCCGCCCCAACCCAGTCGTCCTCTCCATCCCCACTGAC 3190  
Qy 3121 GTCTTTGCCAGCTCCTGTGTCAGAGAAAGGCCCAATTTGCCCGGAAATTCAGGCTTTGCA 3180  
Db 3191 GTCTTTGCCAGCTCCTGTGTCAGAGAAAGGCCCAATTTGCCCGGAAATTCAGGCTTTGCA 3250  
Qy 3181 GGAGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC 3240  
Db 3251 GGAGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC 3310  
Qy 3241 TGGCTACAAATCTGTCTGAGGATCAGCCTCACCCACCGACCATCCCCCTTCAACCTCAT 3300  
Db 3311 TGGCTACAAATCTGTCTGAGGATCAGCCTCACCCACCGACCATCCCCCTTCAACCTCAT 3370  
Qy 3301 GAAGTGCACCTCATGTTAGCGGTGGAGGGCCGCTCTTCAAGAAAGTGTGCTGCAGC 3360  
Db 3371 GAAGTGCACCTCATGTTAGCGGTGGAGGGCCGCTCTTCAAGAAAGTGTGCTGCAGC 3430  
Qy 3361 CCCAGACCTGTCTATTTATTTTCAATTTGGGACAAAGACAGACGTCACAAACAGAGGTGTT 3420  
Db 3431 CCCAGACCTGTCTATTTATTTTCAATTTGGGACAAAGACAGACGTCACAAACAGAGGTGTT 3490  
Qy 3421 TGGGCTTTCAGAAGCCTTTGTTTCCGTGGGTTATGAATATGAATCCTGCCAGATCTAAT 3480  
Db 3491 TGGGCTTTCAGAAGCCTTTGTTTCCGTGGGTTATGAATATGAATCCTGCCAGATCTAAT 3550  
Qy 3481 CCTGTGGGAAAAAGAACACAGTGTGACGGCTATGAAATTTGACGCTTCAAGCTTGG 3540  
Db 3551 CCTGTGGGAAAAAGAACACAGTGTGACGGCTATGAAATTTGATCGTCAAGCTTGG 3610  
Qy 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCTTGACAAAAG 3600  
Db 3611 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCTTGACAAAAG 3670  
Qy 3601 GAATGGGAGAACCCAGTTTGTGTCAGCAGCCTCTCTCATTTGGGAGCATCATGGGCAA 3660  
Db 3671 GAATGGGAGAACCCAGTTTGTGTCAGCAGCCTCTCTCATTTGGGAGCATCATGGGCAA 3730  
Qy 3661 TGGGCGCGGAGAGCATCTCTGCCCCAGCTGCAACGGCCTTGTGACGGGAAACAGCT 3720  
Db 3731 TGGGCGCGGAGAGCATCTCTGCCCCAGCTGCAACGGCCTTGTGACGGGAAACAGCT 3790  
Qy 3721 CCTGGCCCCAGTGGCCCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA 3780

Db 3791 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA 3850  
Qy 3781 CTACATTAGAAGGATCTTCCCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGAATAA 3840  
Db 3851 CTACATTAGAAGGATCTTCCCCCTCTGGAATGTCAACCAACATCCTAGAGCTG----- 3902  
Qy 3841 AGATTTCAAGACATAGTCAAGTCCAGCACAAATACTA CTTGGCCACAGACCCCATGAG 3900  
Db 3903 -----AGTCAAGTCCAGCACAAATACTA CTTGGCCACAGACCCCATGAG 3949  
Qy 3901 TGGGGCCGTCTTCTTCTGACAGCAACAGCCGCGGGTCTTTAAATCAAGTCCACTGT 3960  
Db 3950 TGGGGCCGTCTTCTTCTGACAGCAACAGCCGCGGGTCTTTAAATCAAGTCCACTGT 4009  
Qy 3961 GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAAGTGCCT 4020  
Db 4010 GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAAGTGCCT 4069  
Qy 4021 CCCCTTTGATGACACTCGCTGCGGGATGTTGGGAAGGCCACAGAAAGCCACACTCACCAA 4080  
Db 4070 CCCCTTTGATGACACTCGCTGCGGGATGTTGGGAAGGCCACAGAAAGCCACACTCACCAA 4129  
Qy 4081 TCCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGTGGATGGCACCATGAT 4140  
Db 4130 TCCCAGGGGCATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGTGGATGGCACCATGAT 4189  
Qy 4141 CAGACGCATCGATCAGATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC 4200  
Db 4190 CAGACGCATCGATCAGATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC 4249  
Qy 4201 AGCCCGGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCCAGGTAAGACTGGAGTGGCC 4260  
Db 4250 AGCCCGGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCCAGGTTCACTCGAGTGGCC 4309  
Qy 4261 CACAGACTTAGCCATCAACCCAAATGGACAACTCACTTATGTCCTCGACAAACAATGTT 4320  
Db 4310 CACAGACTTAGCCATCAACCCAAATGGACAACTCACTTATGTCCTCGACAAACAATGTT 4369  
Qy 4321 CTTGCAAAATCTCTGAAAAACCAACAGGTGCGCATTTGTCGCGGAGGCCCATGCACTGCCA 4380  
Db 4370 CTTGCAAAATCTCTGAAAAACCAACAGGTGCGCATTTGTCGCGGAGGCCCATGCACTGCCA 4429  
Qy 4381 GGTCCCTGGCATTTGACCACTTCTGTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440  
Db 4430 GGTCCCTGGCATTTGACCACTTCTGTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4489  
Qy 4441 AGCCACCGCTTTGGCTGTTTACACAATGGGTCCTGTATATTGCTGAGACTGATGAA 4500  
Db 4490 AGCCACCGCTTTGGCTGTTTACACAATGGGTCCTGTATATTGCTGAGACTGATGAA 4549  
Qy 4501 AAAGATCAACCGCATCAGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4560  
Db 4550 AAAGATCAACCGCATCAGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4609  
Qy 4561 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTTCTGGAGACGATGG 4620  
Db 4610 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTTCTGGAGACGATGG 4669  
Qy 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGTGCTGATGGGA 4680  
Db 4670 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGTGATGGGA 4729  
Qy 4681 GCTTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAAACAAGCCTTT 4740  
Db 4730 GCTTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAAACAAGCCTTT 4789  
Qy 4741 CCTCAACACCCAGAACATGTATGAGTGTCTTCAACAATTGACCAGGAGCTCTATCTGTT 4800  
Db 4790 CCTCAACACCCAGAACATGTATGAGTGTCTTCAACAATTGACCAGGAGCTCTATCTGTT 4849  
Qy 4801 TGATACCAACCGGCAAGCCTGTACACCCAAAGCCTGCCACAGGAGACTACTGTACAA 4860  
Db 4850 TGATACCAACCGGCAAGCCTGTACACCCAAAGCCTGCCACAGGAGACTACTGTACAA 4909

QY 4861 CTTCACTTACCTGGGACGGCGACATCACACTCATCACAGACAAATGGCAACATGGT 4920  
Db 4910 CTTCACTTACCTGGGACGGCGACATCACACTCATCACAGACAAATGGCAACATGGT 4969  
QY 4921 AAATGTCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT 4980  
Db 4970 AAATGTCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT 5029  
QY 4981 GTACTGGGTGACCATGGCACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGACACGA 5040  
Db 5030 GTACTGGGTGACCATGGCACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGACACGA 5089  
QY 5041 GTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAA 5100  
Db 5090 GTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAA 5149  
QY 5101 CGGATGGACAAACATTTTATGATACGACAGCTTTGGCCGCTGACAAAATGTGACCTTCCC 5160  
Db 5150 CGGATGGACAAACATTTTATGATACGACAGCTTTGGCCGCTGACAAAATGTGACCTTCCC 5209  
QY 5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAAGTGATGTCAGGTAGA 5220  
Db 5210 TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAAGTGATGTCAGGTAGA 5269  
QY 5221 GACCTCCAGCAAGGATGATGTCAACCAATACCAACCTGTGCTGCTCAGGCGCTTCTA 5280  
Db 5270 GACCTCCAGCAAGGATGATGTCAACCAATACCAACCTGTGCTGCTCAGGCGCTTCTA 5329  
QY 5281 CACACTGCTGCCAACAGCAAGTCCGGAACAGCTACTACATCGGGCCGATGGTCTCTTGGCG 5340  
Db 5330 CACACTGCTGCCAACAGCAAGTCCGGAACAGCTACTACATCGGGCCGATGGTCTCTTGGCG 5389  
QY 5341 GCTGCTGCTGCCAACAGCAAGTCCGGAACAGCTACTACATCGGGCCGATGGTCTCTTGGCG 5400  
Db 5390 GCTGCTGCTGCCAACAGCAAGTCCGGAACAGCTACTACATCGGGCCGATGGTCTCTTGGCG 5449  
QY 5401 CACCGTCAACCCCAACCGTGGGCAAGAGGAATGTCAACCTGCTGCTGCTGCTGCTGCTGCT 5460  
Db 5450 CACCGTCAACCCCAACCGTGGGCAAGAGGAATGTCAACCTGCTGCTGCTGCTGCTGCTGCT 5509  
QY 5461 CCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGCTTCTTGGCG 5520  
Db 5510 CCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGCTTCTTGGCG 5569  
QY 5521 CCGGCTGGGCTGCACAAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCAC 5580  
Db 5570 CCGGCTGGGCTGCACAAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCAC 5629  
QY 5581 AGAGAGATCTATGATGACCAACCGCAAGTTTCAACCTTTCGATTTCTGTACGACCAAGCGGG 5640  
Db 5630 AGAGAGATCTATGATGACCAACCGCAAGTTTCAACCTTTCGATTTCTGTACGACCAAGCGGG 5689  
QY 5641 GCGGCCAGCCTCTGGTCAACCCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC 5700  
Db 5690 GCGGCCAGCCTCTGGTCAACCCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC 5749  
QY 5701 TGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAGAATGGAATACGACCA 5760  
Db 5750 TGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAGAATGGAATACGACCA 5809  
QY 5761 GCGGGCCGATCACATCCAGGATCTTCTGCTGATGGGAAGACATGGAGCTACACATACTT 5820  
Db 5810 GCGGGCCGATCACATCCAGGATCTTCTGCTGATGGGAAGACATGGAGCTACACATACTT 5869  
QY 5821 AGAGAGTCCATGGTGTCTGCTACTACAGCCAGAGGAGTATATCTTTGAGTTCGACAA 5880  
Db 5870 AGAGAGTCCATGGTGTCTGCTACTACAGCCAGAGGAGTATATCTTTGAGTTCGACAA 5929  
QY 5881 GAATGACCGCTCTCTTCTGTGACGATGCCCAACGTGGCGGCGAGACACTAGAGACCAT 5940  
Db 5930 GAATGACCGCTCTCTTCTGTGACGATGCCCAACGTGGCGGCGAGACACTAGAGACCAT 5989

QY 5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGGCTCAGTCAT 6000  
Db 5990 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGGCTCAGTCAT 6049  
QY 6001 ACAGGACTTCACTGAGGATGGCACCTCCTTACACACTTCTTACCTGGGCACTGGCCGCGAG 6060  
Db 6050 ACAGGACTTCACTGAGGATGGCACCTCCTTACACACTTCTTACCTGGGCACTGGCCGCGAG 6109  
QY 6061 GGTGATATACAAGTATGGCAAACTGTCAAAGTGGCAGAGACGCTCTATGACACCAACAA 6120  
Db 6110 GGTGATATACAAGTATGGCAAACTGTCAAAGTGGCAGAGACGCTCTATGACACCAACAA 6169  
QY 6121 GGTGAGTTTCACTATGACGAGACGGCAGGATGTGAAGACCATCAACCTACAGAATGA 6180  
Db 6170 GGTGAGTTTCACTATGACGAGACGGCAGGATGTGAAGACCATCAACCTACAGAATGA 6229  
QY 6181 GGGCTTCACTGCACCATCCGCTACCGTACAGATTGGGCCCTGATTGACCGACAGATCTT 6240  
Db 6230 GGGCTTCACTGCACCATCCGCTACCGTACAGATTGGGCCCTGATTGACCGACAGATCTT 6289  
QY 6241 CCGTTTCACTGAGGAAGGCATGGTCAACGCCCGTTTGTACTACAACCTATGACAAACAGCTT 6300  
Db 6290 CCGTTTCACTGAGGAAGGCATGGTCAACGCCCGTTTGTACTACAACCTATGACAAACAGCTT 6349  
QY 6301 CCGGTTGACCAAGCATGCAGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCG 6360  
Db 6350 CCGGTTGACCAAGCATGCAGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCG 6409  
QY 6361 CTATGATGATGTCTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTATTACTATGA 6420  
Db 6410 CTATGATGATGTCTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTATTACTATGA 6469  
QY 6421 CATTAAACAGATCATCACACAGCTGTGATGACCCCAACCAAGCAATTTTGTGATGATATGG 6480  
Db 6470 CATTAAACAGATCATCACACAGCTGTGATGACCCCAACCAAGCAATTTTGTGATGATATGG 6529  
QY 5481 CAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAAGTGGATGACCGTCCA 6540  
Db 6530 CAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAAGTGGATGACCGTCCA 6589  
QY 6541 GTATGATAACATGGGCGAGTATGAGTGAAGAGGAGCTGAAGTAGGACCTACGCCAATAC 6600  
Db 6590 GTATGATAACATGGGCGAGTATGAGTGAAGAGGAGCTGAAGTAGGACCTACGCCAATAC 6649  
QY 6601 CACTCGTACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6660  
Db 6650 CACTCGTACTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6709  
QY 6661 CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGACCTTACTGAGCCCTGG 6720  
Db 6710 CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGACCTTACTGAGCCCTGG 6769  
QY 6721 GAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGG 6780  
Db 6770 GAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGG 6829  
QY 6781 TGACGTGCAATACAAGATGGATGAGGATGGCTTCTTCTGAGGACGCGGGCGGTGATATCTT 6840  
Db 6830 TGACGTGCAATACAAGATGGATGAGGATGGCTTCTTCTGAGGACGCGGGCGGTGATATCTT 6889  
QY 6841 TGAGTACAACCTCAGCTGGCCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGT 6900  
Db 6890 TGAGTACAACCTCAGCTGGCCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGT 6949  
QY 6901 CAGGTACCGCTACGATGGCCTGGGCGCGCGCTGTCCAGCAAGACAGCCACCA 6960  
Db 6950 CAGGTACCGCTACGATGGCCTGGGCGCGCGCTGTCCAGCAAGACAGCCACCA 7009  
QY 6961 CCTGAGTCTTCTATGACAGCTGACCAACCCCAACCAAGGTACCCACCTGTACAACCA 7020  
Db 7010 CCTGAGTCTTCTATGACAGCTGACCAACCCCAACCAAGGTACCCACCTGTACAACCA 7069  
QY 7021 CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGSACACCTCTTTGCCATGGA 7080

Db 7070 CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGAAGGACACCTCTTTGCCATGGA 7129  
QY 7081 GCTGAGCAGTGGTGATGAGTTTACATAGCTTGTGACAAACATCGGAGCCCTCTTTGCTGT 7140  
Db 7130 GCTGAGCAGTGGTGATGAGTTTACATAGCTTGTGACAAACATCGGAGCCCTCTTTGCTGT 7189  
QY 7141 CTTTAGTGAACAGGTTTGATGATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTA 7200  
Db 7190 CTTTAGTGAACAGGTTTGATGATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTA 7249  
QY 7201 CATGGATACCAACCCCAACTTTTCAGATCATATFAGGCTACCATGGTGCCCTCTATGATCC 7260  
Db 7250 CATGGATACCAACCCCAACTTTTCAGATCATATFAGGCTACCATGGTGCCCTCTATGATCC 7309  
QY 7261 ACTCACCAAGCTTGTCCACATGGCGCGGAGATATGATGTGCTGGCGGACGCTGGAC 7320  
Db 7310 ACTCACCAAGCTTGTCCACATGGCGCGGAGATATGATGTGCTGGCGGACGCTGGAC 7369  
QY 7321 TAGCCAGACCCAGCTGTGTGAAGCACCTTTAGTAGCAGCAACGTCATGCTCTTTTAATCT 7380  
Db 7370 TAGCCAGACCCAGCTGTGTGAAGCACCTTTAGTAGCAGCAACGTCATGCTCTTTTAATCT 7429  
QY 7381 CTATATGTTCAAAAACAAACCCCATCAGCAACTCCAGGACATCAAGTGTTCATGAC 7440  
Db 7430 CTATATGTTCAAAAACAAACCCCATCAGCAACTCCAGGACATCAAGTGTTCATGAC 7489  
QY 7441 AGATGTTAACAGCTGGCTGCTCACCTTTTGATTCCAGCTACACAACTGATCCCTGGTTA 7500  
Db 7490 AGATGTTAACAGCTGGCTGCTCACCTTTTGATTCCAGCTACACAACTGATCCCTGGTTA 7549  
QY 7501 TCCCAACCCAGACATGGATGCCATGGAAACCTCTACGAGCTCATCCACACACAGATGAA 7560  
Db 7550 TCCCAACCCAGACATGGATGCCATGGAAACCTCTACGAGCTCATCCACACACAGATGAA 7609  
QY 7561 AACCCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA 7620  
Db 7610 AACCCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA 7669  
QY 7621 GGTCAAGGCCTTTGTCACTTAGAACGGTTTGACAGCTCTATGGCTCCACAAATCACCAG 7680  
Db 7670 GGTCAAGGCCTTTGTCACTTAGAACGGTTTGACAGCTCTATGGCTCCACAAATCACCAG 7729  
QY 7681 CTGCCAGCAGGCTCCAAAGAACCAAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGG 7740  
Db 7730 CTGCCAGCAGGCTCCAAAGAACCAAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGG 7789  
QY 7741 GGTCAAGTTTGCCTTGAAGATGGCCGAGTGACCAAGACATCATCAGTGTGGCCAAATGA 7800  
Db 7790 GGTCAAGTTTGCCTTGAAGATGGCCGAGTGACCAAGACATCATCAGTGTGGCCAAATGA 7849  
QY 7801 GGATGGCGGAAGGTTGCTGCACTTTGAACCAATGCCACTACCTAGAGAACCTGCACCT 7860  
Db 7850 GGATGGCGGAAGGTTGCTGCACTTTGAACCAATGCCACTACCTAGAGAACCTGCACCT 7909  
QY 7861 CACCAATTGATGGGGTGATACCCATTACTTTGTGAACCAAGGACCTTCAGAGGTTGACCT 7920  
Db 7910 CACCAATTGATGGGGTGATACCCATTACTTTGTGAACCAAGGACCTTCAGAGGTTGACCT 7969  
QY 7921 GGCCATCCTGGGCTCAGTGGGGCGGCGAACCTTGAGAAATGGGTCAACGTCACCTGT 7980  
Db 7970 GGCCATCCTGGGCTCAGTGGGGCGGCGAACCTTGAGAAATGGGTCAACGTCACCTGT 8029  
QY 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8040  
Db 8030 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8089  
QY 8041 GTACGGGCACTGTGCTTGAACACACAGCTACGGGACAACTTGGATGAGGAGAGGCACG 8100  
Db 8090 GTACGGGCACTGTGCTTGAACACACAGCTACGGGACAACTTGGATGAGGAGAGGCACG 8149  
QY 8101 GGTCTGAGCTGGCCCGGACAGAGCCGCTGCGCCAAAGCTGGGCCCGGACGACGAGAG 8160

8150 GGTCTCTGAGCTGGCCCGGACAGAGCCGTGCGCCCAAGCGTGGGCCCGGACGAGAG 8209  
QY 8161 ACTGCGGAAGGGAGGAAGGCCTGCGGCTGGACAGAGGGGAGAGAGCAGAGTCT 8220  
Db 8210 ACTGCGGAAGGGAGGAAGGCCTGCGGCTGGACAGAGGGGAGAGAGCAGAGTCT 8269  
QY 8221 GAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTCTGTGATCTCTCTGAGCAGTACCC 8280  
Db 8270 GAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTCTGTGATCTCTCTGAGCAGTACCC 8329  
QY 8281 AGAACTGTGACAGCGGCCCAACAACATCCACTTCATGAGACAGAGCGAGATGGCCGGAG 8340  
Db 8330 AGAACTGTGACAGCGGCCCAACAACATCCACTTCATGAGACAGAGCGAGATGGCCGGAG 8389  
QY 8341 GTGACAGAGAGGAC 8354  
Db 8390 GTGACAGAGAGGAC 8403

RESULT 3  
ABN85378  
ID ABN85378 standard; DNA; 8438 BP.  
XX AC ABN85378;  
XX AC ABN85378;  
DT 21-OCT-2002 (first entry)  
XX Human NOV1, TEN-M4 like protein, coding sequence.  
DE Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;  
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;  
KW Antiasthmatic; Nephrotropic; Hepatotrophic; Neuroprotective; Nootropic;  
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;  
KW Gene Therapy; NOV; cancer; heart disease; inflammation;  
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;  
KW asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease;  
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;  
KW TEN-M4 like protein; chromosome 11; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT CDS 4. 8395  
FT /\*tag= a  
FT /trans\_except= (pos: 1138. 1147,aa:Met)  
FT /product= "NOV1 protein"  
XX  
PN WO200255704-A2.  
XX  
PD 18-JUL-2002.  
XX  
PF 09-JAN-2002; 2002WO-US0000554.  
XX  
PR 09-JAN-2001; 2001US-0260417P.  
PR 10-JAN-2001; 2001US-0260831P.  
PR 28-FEB-2001; 2001US-0272338P.  
PR 09-MAR-2001; 2001US-0274876P.  
PR 18-APR-2001; 2001US-0284704P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;  
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;  
PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;  
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;  
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;  
PI Gunther E, Stone DJ;  
XX  
DR WPI; 2002-590674/63.  
DR P-PSDB; ABB98401.  
XX  
PT NOVX polypeptides and encoding polynucleotides, useful for preventing or  
treating NOVX-associated disorders e.g. cancer, inflammation, or



Db 1852 CTGGAAGGCAACCGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGG 1911  
Qy 1945 CCATGGCACCTGCATCACGGGCACCTGCATCTGCAACCCCTGGCTACAAAGGCGAGAGCTG 2004  
Db 1912 TCGTGGGATTTGTATCATGTGGGCTCCTGTGCTTGCAACTCAGGATACAAAGGAGAAAGTTG 1971  
Qy 2005 TGAGGAAGTGGACTGCATGGAACCCACACATGTTTCAGGCGGGGTGTCTGCGTGAGAGCGGA 2064  
Db 1972 TGAAGAAGCTGACTGTATAGACCCCTGGGTGTTCTTAATCATGTGTGTAATCTGAAGACCATGTGTCC 2031  
Qy 2065 ATGCCATTGCTTTGTGGGATGGGAGGCACCAACTGCGAGACCCCGGCTTTGACGCTGTGACCC 2124  
Db 2032 ATGTCACTGCAGTCCAGGATGGGAGGTAGCAATTGTGAAATACTGAAGACCATGTGTCC 2091  
Qy 2125 AGACCAGTGTTCAGGCCACGGAACCTTCCTCCCGACACCGGGCTTTGACGCTGTGACCC 2184  
Db 2092 AGACCAGTGTCCGGCCACGGAACGTAFTCTTCAAGAAAGTGGCTCCTGCACGTGTGACCC 2151  
Qy 2185 AACCTGGACTGGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGCTGGCCATGG 2244  
Db 2152 TAACTGGACTGGCCAGACTGCTCAACGAAATATGTTCTGTGACTGTGGCTCACACGG 2211  
Qy 2245 CGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGGACGCTGCGACCA 2304  
Db 2212 CGTTTGATGGGGGGACGTGTGCTGTGAAGAAGGCTGACGGGCCACGCTGTAATCA 2271  
Qy 2305 GCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACCTGCCGACGGCAAGTGGAGTG 2364  
Db 2272 GAGAGCCTGCCACCCCGCTGTGCCGAGCACGGGACCTGCCGACGGCAAGTGGAGTG 2331  
Qy 2365 CAGCCCTGGCTGGAATGGCAACACTGCACCATC-----GCTCACTATCTGGATAGGGT 2418  
Db 2332 CAGCCCTGGCTGGAATGGCAACACTGCACCATCTCCCTAGCTCACTATCTGGATAGGGT 2391  
Qy 2419 AGTTAA-----AGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGA 2472  
Db 2392 AGTTAAACTTTTCAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGA 2451  
Qy 2473 CTTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTC 2532  
Db 2452 CCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTC 2511  
Qy 2533 CATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTGTGGACTGCAT 2592  
Db 2512 CATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTGTGGACTGCAT 2571  
Qy 2593 GGACCTTGACTGTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCCCTGGCTCCCC 2652  
Db 2572 GGACCTTGACTGTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCCCTGGCTCCCC 2631  
Qy 2653 TAACCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTACAGCAGAACCTTACACTC 2712  
Db 2632 TAACCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTACAGCAGAACCTTACACTC 2691  
Qy 2713 CTTCTATGACCGCATCAAGTTCCTCGTGGCAGGACAGCAGCATATATCCCGGGA 2772  
Db 2692 CTTCTATGACCGCATCAAGTTCCTCGTGGCAGGACAGCAGCATATATCCCGGGA 2751  
Qy 2773 GAACCCCTTTGATGGAGGGGATGCTTGTGTTATTCTGGCCCAAGTGATGACATCAGATGG 2832  
Db 2752 GAACCCCTTTGATGGAGGGGATGCTTGTGTTATTCTGGCCCAAGTGATGACATCAGATGG 2811  
Qy 2833 AACCCCTCTGGTTGGTGAACATCAGTTTTTGTCAATAACCTCTCTTTGGATATACAAT 2892  
Db 2812 AACCCCTCTGGTTGGTGAACATCAGTTTTTGTCAATAACCTCTCTTTGGATATACAAT 2871  
Qy 2893 CAGAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGGGCATCTCCATCATCTCGG 2952  
Db 2872 CAGAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGGGCATCTCCATCATCTCGG 2931  
Qy 2953 GTTCAGCGGGCACCTTTTCATACACAGGAGCACCCCTGTGGCTGCCATCGCTT 3012

Db 2932 GTTCGAGCGGGCACCTTTTCATCACAGGAGCACACCTGTGTGGCTGCCATGGGATCGCTT 2991  
Qy 3013 CTTTGTTCATGGAACCATCATCATGAGACATGAGGAGAAATGAGATTCAGCTGTGACCT 3072  
Db 2992 CTTTGTTCATGGAACCATCATCATGAGACATGAGGAGAAATGAGATTCAGCTGTGACCT 3051  
Qy 3073 GAGCAATTTTCCCGCCCAACCCAGTCTCTCTCCATCCCACTGAGCTCCTTCGCCAG 3132  
Db 3052 GAGCAATTTTCCCGCCCAACCCAGTCTCTCTCCATCCCACTGAGCTCCTTCGCCAG 3111  
Qy 3133 CTCCTGTGCAGAGAAAGGCCCTTGTGCCGAAAATTCAGGCTTTGACAGGAGAAATCTC 3192  
Db 3112 CTCCTGTGCAGAGAAAGGCCCTTGTGCCGAAAATTCAGGCTTTGACAGGAGAAATCTC 3171  
Qy 3193 TATCTCTGGCTGCAAGATGAGGCTGAGCTACTGAGCAGCCGACCCCTGGCTACAAATC 3252  
Db 3172 TATCTCTGGCTGCAAGATGAGGCTGAGCTACTGAGCAGCCGACCCCTGGCTACAAATC 3231  
Qy 3253 TGTCTGAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCT 3312  
Db 3232 TGTCTGAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCT 3291  
Qy 3313 CATGGTAGCGGTGGAGCGCCCTCTTTCAGGAAGTGGTTGCTGCAGCCCCCAGACCTGTC 3372  
Db 3292 CATGGTAGCGGTGGAGCGCCCTCTTTCAGGAAGTGGTTGCTGCAGCCCCCAGACCTGTC 3351  
Qy 3373 CTATTATTTTCATTTGGACAAGACAGACGCTCTACAAACAGAGGTGTTGGGCTTTTCAGA 3432  
Db 3352 CTATTATTTTCATTTGGACAAGACAGACGCTCTACAAACAGAGGTGTTGGGCTTTTCAGA 3411  
Qy 3433 AGCCTTTGTTTCCGTGGGTTATGAATATGAATTCCTGCCAGATCTAATCTGTGGGAAAA 3492  
Db 3412 AGCCTTTGTTTCCGTGGGTTATGAATATGAATTCCTGCCAGATCTAATCTGTGGGAAAA 3471  
Qy 3493 AAGAAACAACAGTGTGCAGGGCTATGAAATTGACGCGTCCAAGCTTGGAGGATGGAGCCT 3552  
Db 3472 AAGAAACAACAGTGTGCAGGGCTATGAAATTGACGCGTCCAAGCTTGGAGGATGGAGCCT 3531  
Qy 3553 AGACAAAACATCATGCCCTCAACATTTCAAAGTGGT---ATCCTGCACAAAGGGAATGGGA 3609  
Db 3532 AGACAAAACATCATGCCCTCAACATTTCAAAGTGGTGGCATCTCTGCACAAAGGGAATGGGA 3591  
Qy 3610 GAACCAAGTTTGTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATCATGGCAATGGCGCCG 3669  
Db 3592 GAACCAAGTTTGTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATCATGGCAATGGCGCCG 3651  
Qy 3670 GAGAAAGCATCTCCTGCCCAAGCTGCAACGGCCCTGTGACGGCAACAAGCTCCTGGCCCC 3729  
Db 3652 GAGAAAGCATCTCCTGCCCAAGCTGCAACGGCCCTGTGACGGCAACAAGCTCCTGGCCCC 3711  
Qy 3730 AGTGGCCCTCACCTGTGCTGTGACGGGAGCCTCTATGTGGTGATTTCAACTACATTAG 3789  
Db 3712 AGTGGCCCTCACCTGTGCTGTGACGGGAGCCTCTATGTGGTGATTTCAACTACATTAG 3771  
Qy 3790 AAGGATCTTCCCTCTGGAATGTCAACCAACATCTCTAGAGCTGAGG-----AATAAAGA 3843  
Db 3772 AAGGATCTTCCCTCTGGAATGTCAACCAACATCTCTAGAGCTGAGGGTCAGAAATAAGA 3831  
Qy 3844 TTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGG 3903  
Db 3832 TTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGG 3891  
Qy 3904 GGCGCTCTTCCCTTCTGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGTGGT 3963  
Db 3892 GGCGCTCTTCCCTTCTGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGTGGT 3951  
Qy 3964 GGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTGCGGGGACAGGTGACAGTGCCTCCC 4023  
Db 3952 GGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTGCGGGGACAGGTGACAGTGCCTCCC 4011  
Qy 4024 CTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCC 4083  
Db 4012 CTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCC 4071

QY	4084	CAGGGGT-----ATTACAGTGGACAAGTTTGGGCTGATCTACTTTCGTGGATGGCAC	4134
Db	4072	CAGGGGTCCCCCAGGCATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCAC	4131
QY	4135	CATGATCAGACGCATCGATCAGAATGGATCATCTCCACCCCTGCTCGGCTCTAATGATCT	4194
Db	4132	CATGATCAGACGCATCGATCAGAATGGATCATCTCCACCCCTGCTCGGCTCTAATGATCT	4191
QY	4195	CACATCAGCCCGCCCACTCAGCTGTGATCTGTCTCATGGATATTCCCAGGTAAGA----	4249
Db	4192	CACATCAGCCCGCCCACTCAGCTGTGATCTGTCTCATGGATATTCCCAGGTAAGACAGT	4251
QY	4250	----CTGGAGTGGCCCAACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCT	4305
Db	4252	TCACCTGGAGTGGCCCAACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCT	4311
QY	4306	CGACAACAAATGTGGTCTCTGCAAAATCTCTGAACCAACCAAGTGGCATTTGCGCCGGAG	4365
Db	4312	CGACAACAAATGTGGTCTCTGCAAAATCTCTGAACCAACCAAGTGGCATTTGCGCCGGAG	4371
QY	4366	GCCCATGCACTGCCAGGTCCCTGGCAFTGACCACTTCCTGCTAAGCAAGGTGGCCATCCA	4425
Db	4372	GCCCATGCACTGCCAGGTCCCTGGCAFTGACCACTTCCTGCTAAGCAAGGTGGCCATCCA	4431
QY	4426	CGCAACCTGGAGTCAGCCACCGCTTTGGCTGTTTTACACAATGGGTCTCTGTATATTGC	4485
Db	4432	CGCAACCTGGAGTCAGCCACCGCTTTGGCTGTTTTACACAATGGGTCTCTGTATATTGC	4491
QY	4486	TGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTC	4545
Db	4492	TGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTC	4551
QY	4546	ACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTT	4605
Db	4552	ACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTT	4611
QY	4606	TTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGT	4665
Db	4612	TTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGT	4671
QY	4666	GTGTCTGATGGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCG	4725
Db	4672	GTGTCTGATGGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCG	4731
QY	4726	GAAGAACAAAGCCTTTCTCAACACCCAGAACATGATGAGTGTCTTCAACCAATTGACCA	4785
Db	4732	GAAGAACAAAGCCTTTCTCAACACCCAGAACATGATGAGTGTCTTCAACCAATTGACCA	4791
QY	4786	GGAGTCTATCTGTTTGATACCAACCGGAAGCACCTGTACACCCAAAGCCTGCCACAGG	4845
Db	4792	GGAGTCTATCTGTTTGATACCAACCGGAAGCACCTGTACACCCAAAGCCTGCCACAGG	4851
QY	4846	AGACTACCTGTACAACTTCACCTACACTGGGACGGCGACATCACACTCATCACAGACAA	4905
Db	4852	AGACTACCTGTACAACTTCACCTACACTGGGACGGCGACATCACACTCATCACAGACAA	4911
QY	4906	CAATGGCAACATGGTAAATGTCGCCCGAGACTCTACTGGGATGCCCCCTCTGGCTGGTGT	4965
Db	4912	CAATGGCAACATGGTAAATGTCGCCCGAGACTCTACTGGGATGCCCCCTCTGGCTGGTGT	4971
QY	4966	CCCAGATGGCCAGGTGTAAGTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGAC	5025
Db	4972	CCCAGATGGCCAGGTGTAAGTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGAC	5031
QY	5026	CACACAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAAC	5085
Db	5032	CACACAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAAC	5091
QY	5086	CAAAAGCAATGAAAACCGGATGGACAAACATTTTATGATGACGACAGCTTTGGCCGCTGAC	5145
Db	5092	CAAAAGCAATGAAAACCGGATGGACAAACATTTTATGATGACGACAGCTTTGGCCGCTGAC	5151

QY	5146	AAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGT	5205
Db	5152	AAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGT	5211
QY	5206	GCATGTCCAGGTAGAGACCTCCAGCAAGSATGATGTCAACATAACCAACCAACCTGTCTGC	5265
Db	5212	GCATGTCCAGGTAGAGACCTCCAGCAAGSATGATGTCAACATAACCAACCAACCTGTCTGC	5271
QY	5266	CTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGC	5325
Db	5272	CTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGC	5331
QY	5326	CGATGGCTCCTTGGGCTGCTGTGGCCAAACGGCATGGAGGTGGGCTGCAGACTGAGCC	5385
Db	5332	CGATGGCTCCTTGGGCTGCTGTGGCCAAACGGCATGGAGGTGGGCTGCAGACTGAGCC	5391
QY	5386	CCACTTGTGTGGTGGCACCGTCAACCCACCGTGGGCAAGAGGAATGTCAACGCTGCCCAT	5445
Db	5392	CCACTTGTGTGGTGGCACCGTCAACCCACCGTGGGCAAGAGGAATGTCAACGCTGCCCAT	5451
QY	5446	CGACAAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAAAGAGCAGGCTCGGGGCCAGGT	5505
Db	5452	CGACAAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAAAGAGCAGGCTCGGGGCCAGGT	5511
QY	5506	CACGTCTTTTGGGCGCGGCTGCGGGTG-----CACAAACGAAATCTCCTATCTCT	5556
Db	5512	CACGTCTTTTGGGCGCGGCTGCGGGTGCTCCAGGTTTCAACACCGAAATCTCCTATCTCT	5571
QY	5557	GGACTTTGATCGGTAACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTTCAACCT	5616
Db	5572	GGACTTTGATCGGTAACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTTCAACCT	5631
QY	5617	TCGGATTCTGTACGACACAGGCGGGGCGGCCAGCCCTCTGTTCCACGACGACGAGGTGAA	5676
Db	5632	TCGGATTCTGTACGACACAGGCGGGGCGGCCAGCCCTCTGTTCCACGACGACGAGGTGAA	5691
QY	5677	TGGTGTCAACGTGACATACTCCCTTGGGGTTTACATTGCTGGCATCCAGAGGGGCATCAT	5736
Db	5692	TGGTGTCAACGTGACATACTCCCTTGGGGTTTACATTGCTGGCATCCAGAGGGGCATCAT	5751
QY	5737	GTCTGAAAGAAATGGAATACGACCCAGCGGGCCGCATCACATCCAGGATCTTCGCTGATGG	5796
Db	5752	GTCTGAAAGAAATGGAATACGACCCAGCGGGCCGCATCACATCCAGGATCTTCGCTGATGG	5811
QY	5797	GAAGACATGGAGCTACACATACTTAGAGA-----AGTCCATGGTGTGCTACT	5844
Db	5812	GAAGACATGGAGCTACACATACTTAGAGAAGGCAGGTGTCCAGTCCATGGTGTGCTACT	5871
QY	5845	ACACAGCCAGAGGCAGTATATCTTTGAGTTCGACAAGAAATGACCGCTCTCTTCTGTGAC	5904
Db	5872	ACACAGCCAGAGGCAGTATATCTTTGAGTTCGACAAGAAATGACCGCTCTCTTCTGTGAC	5931
QY	5905	GATGCCCCAACGTGGCGGCAGACACTAGAGACCAATCCGCTCAGTGGGCTACTACAGAAA	5964
Db	5932	GATGCCCCAACGTGGCGGCAGACACTAGAGACCAATCCGCTCAGTGGGCTACTACAGAAA	5991
QY	5965	CATCTATCAGCCCCCTGAGGGCAATGCTCAGTCAATACAGGACTTCACTGAGGATGGGCA	6024
Db	5992	CATCTATCAGCCCCCTGAGGGCAATGCTCAGTCAATACAGGACTTCACTGAGGATGGGCA	6051
QY	6025	CCTCCTTACACACTTCTACCTGGGCACTGGCGCAGGTTGATATACAAGTATGGCAAACT	6084
Db	6052	CCTCCTTACACACTTCTACCTGGGCACTGGCGCAGGTTGATATACAAGTATGGCAAACT	6111
QY	6085	GTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCACTTTTACCTATGACGAGAC	6144
Db	6112	GTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCACTTTTACCTATGACGAGAC	6171
QY	6145	GGCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGGCTTACCTGCACCATCCGCTA	6204
Db	6172	GGCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGGCTTACCTGCACCATCCGCTA	6231
QY	6205	CCGTGAGATTGGGCCCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGT	6264









Db 5035 TTGTTTACTTACCATGGCAATAGTGGCCCTTTAGCCACTAAAGTGATGAACCTGGATGG 5094  
Qy 5108 ACAACATTTTATGATGACGACAGCTTTGGCCGCTTGACAAATGAGCTTCCCTACTGGC 5167  
Db 5095 ACAACGTTTTTTGACTATGACAGTGAAGGTCGTCTGACAAATGTTACGTTTCCAACTGGA 5154  
Qy 5168 CAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAAGTGCATGTCAGGCTAGACACCTCC 5227  
Db 5155 GTGGTCAAAAACCTGCATGGGGACATGGACAAAGGCTATCAAGTGGACATGAGTCACT 5214  
Qy 5228 AGC---AAGGATGATGTACCCATAACCAACCACTGTCTGCTCAGGGCCCTTCTACACA 5284  
Db 5215 AGCCGAGAAGATGTACGATCACTTCAAAATCTGCTCGATCGATCTTTCTTCTACACC 5274  
Qy 5285 CTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGATGGCTCCTTGGGGCTG 5344  
Db 5275 ATGGTTCAAGATCAGTTAAGAAACAGCTACAGATTGGTTATGACGGCTCCCTCAGAAAT 5334  
Qy 5345 CTGCTGGCCAAACGGCATGGAGGTGGCGTGCAGACTGAGCCCCACTTGTGCTGGCCACC 5404  
Db 5335 ATCTACGCCAGTGGCTGGACTCACACTACCAACAGAGCCGACGTTCTGGCTGGCCACC 5394  
Qy 5405 GTCAACCCCAACCGTGGCAAGAGGAATGTACGCTGCCATCGACACAGGCTCAACCTG 5464  
Db 5395 GCTAATCCGACGGTTGCCAAAAGAAACATGACTTTCCTGGCGAGAACGGTCAAAACTTG 5454  
Qy 5465 GTGGAGTGGCCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGTCTTTGGCGCCGG 5524  
Db 5455 GTGGAATGGAGATTCGGAAGAGCAAGCCCAAGGAAAGTCAATGTCTTTGGCCGCAAG 5514  
Qy 5525 CTGCGGGTGCAACAACCGAAATCTCCTATCTCTGACTTTTGATCGCGTAACACGCACAGAG 5584  
Db 5515 CTCAGGGTTAATGGCAGAAACCTCTTTTCAGTTGACTTTGATCGAAACAAAGACAGAA 5574  
Qy 5585 AAGATCTATGATGACCAACCGCAAGTTCAACCTTCGGATTCTGTACGACCAAGCGGGCGG 5644  
Db 5575 AAGATCTATGACGACCAACCGTAAATTTCTACTGAGGATCGCCTACGACAGCTCTGGGCAC 5634  
Qy 5645 CCCAGCCTCTGGTCAACCGACGAGCGGTGAATGGTGTCAACGTGACATACTCCCTGGG 5704  
Db 5635 CCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGGCCGTCAATGTCACTATTATCCACA 5694  
Qy 5705 GGTACATTGTGGATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCAAGCGG 5764  
Db 5695 GGTCAAAATGCCAGCATCCAGCGAGGCCACTAGCGAGAAAGTAGATTATGACGGACAG 5754  
Qy 5765 GGCCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATACTTAGAG 5824  
Db 5755 GGGAGGATCGTGTCTCGGGTCTTTGCTGATGGTAAACATGGAGTTACACATATTTAGAA 5814  
Qy 5825 AAGTCCATGGTGTCTACTACAGCCAGAGGAGTATATCTTTGAGTTCGACAAAGAAAT 5884  
Db 5815 AAGTCCATGGTGTCTCTGCTTCATAGCCAGCGGAGTACATCTTCGAATACGATATGTGG 5874  
Qy 5885 GACCGCCTCTCTTCTGTACGATGCCCAACGTTGGCGGCGAGACACTAGAGACCATCCGC 5944  
Db 5875 GACCGCCTGTCTGCCATCACCATGCCAGTCCCAGTGTGGTCCGACACACCATGCAGACCATCCGA 5934  
Qy 5945 TCAGTGGGCTACTACAGAAACATCTATCAGCCCTCTGAGGCAATGCTCAGTACATACAG 6004  
Db 5935 TCCATTGGCTACTACCGCAACATATACAACCCCGGAAAGCAACGCTCCATCATCACG 5994  
Qy 6005 CACTTCACTGAGGATGGGACCTCTTCAACCTTCTACACCTTCTACCTGGGCACTGGCCGAGGGTG 6064  
Db 5995 GACTACAACGGAAGGGTGTCTTCAAAACAGCTTTCTTGGGTACAAGTCCGAGGGTC 6054  
Qy 6065 ATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGAGCTCTATGACACCAACCAAGGTC 6124  
Db 6055 TTATTCAAATACAGAAGGAGAGCTAGGCTCTCAGAAATTTTATATGATACACAAAGATC 6114  
Qy 6125 AGTTTCACTATGACGAGCGGACGATCTGTGAGACCATCAACCTACAGAATGAGGGC 6184

Db 6115 AGTTTACCTATGATGAAACAGCAGGAGTCTCTAAAGACAGTAACCTCCAGAGTGATGGT 6174  
Qy 6185 TTCACCTGCACCATCCGCTACCGTACAGATTGGGCCCTCTGATTGACGACAGATCTTCCGC 6244  
Db 6175 TTTATTTGCACCATTAGATACAGGCAAAATGGTCCCTGATTGACAGGACAGATTTCCGC 6234  
Qy 6245 TTCACGTAGGAAGGCATGGTCAACGCCCTTTTGACTACAACTATGACAAACAGCTTCCGG 6304  
Db 6235 TTTAGTGAAGATGGGATGGTAAATGCAAGATTTGACTATAGCTATGACAAACAGCTTTCGA 6294  
Qy 6305 GTGACCAGCATGCAGGCTGTGATCAACGAGACCCCACTGCCAATGATCTCTATCGCTAT 6364  
Db 6295 GTGACCAGCATGCAGGCTGTGATCAATGAACGCCCACTGCCAATGATCTGTATCAGTTT 6354  
Qy 6365 GATGATGTGTCAAGCAAGACAGAGCAGTGTGGGAAAGTTTGGTGTCAATTTACTATGACATT 6424  
Db 6355 GATGACATTTCTGCCAAAGTTGAGCAGTTTGGAAAGTTTGGAGTTATATATATGATATT 6414  
Qy 6425 AACCAGATCATCAACACAGCTGTCTATGACCCACACCAAGCATTTTGTATGATATGGCAGG 6484  
Db 6415 AACCAGATCATTTCTACAGCTGTAATGACCTATACGAAGCACTTTGATGCTCATGGCCGT 6474  
Qy 6485 ATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTAT 6544  
Db 6475 ATCAAGGAGATTCAATATGAGATATTTCAGTCTGCTCATGTACTGGATTACAATTCAGTAT 6534  
Qy 6545 GATAACATGGGGCAGTAGTGAAGAGGAGCTGAAGGTAGGACCTACGCCCAATACCACT 6604  
Db 6535 GATAACATGGGTCCGGTAACCAAGAGAGAGATTAAATAGGGCCCTTTGCCAACACCACC 6594  
Qy 6605 CGTACTCTATGATGATGCTGACGCGCAGCTGCAGACAGTCTCCATCAATGACAAG 6664  
Db 6595 AAATATGCTTATGAATATGATGTTGATGGACAGCTCCAAACAGTTTACCTCAATGAAAAG 6654  
Qy 6665 CCCTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAAC 6724  
Db 6655 ATAATGTGGCGGTACAACTACGATCTGAATGGAAACCTCCATTTACTGAACCCAAAGTAAC 6714  
Qy 6725 AGTGACGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGAC 6784  
Db 6715 AGTGCGCTCTGACACCCCTTCGCTATGACCTGCGAGACAGAAATCACTCGACTGGGTGAT 6774  
Qy 6785 GTGCAATACAAGATGGATGAGGATGGCTTCCTGAGGCGAGCGGGCGGTGATATCTTTGAG 6844  
Db 6775 GTTCAATATCGGTTGGATGAAGATGGTTTCTACGTCAAAGGGGACGGAATCTTTTGA 6834  
Qy 6845 TACAACTCAGCTGGCTGCTCATCAAGGCTTACAAACCGGCTGGAGCTGGAGTGTCAAG 6904  
Db 6835 TATAGCTCCAAGGGCTTCTAACTCGAGTTTACAGTAAAGGCGAGTGGCTGGACAGTATC 6894  
Qy 6905 TACCGCTACGATGGCTGGGGCGGCGGTGCTGTCAGCAAGAGCGACCAAGCCACCTG 6964  
Db 6895 TACCGTTATGACGCGCTGGGAAGGCGTGTCTTAGCAAAACCAAGTCTTAGGACAGCACCTG 6954  
Qy 6965 CAGTTCTTCTATGACAGCTGACCAACCCCAAGGTCACCCACCTGTACAAACCACTCC 7024  
Db 6955 CAGTTTTTTTATGCTGACTTAACCTTATCCCACTAGGATTTACTCATGTCTACAAACCATTCG 7014  
Qy 7025 AGCTCTGAGATCACTCCCTCTACTACGACTTGAAGGACACCTCTTTTGGCATGGAGCTG 7084  
Db 7015 AGTTCAGAAATTACCTCCCTGTATTATGATCTCCAAGGACATCTTTTGGCATGGAAATC 7074  
Qy 7085 ACCAGTGGTGTGATGAGTTTACATAGCTTGTGACAAACATCGGACCCCTCTTGTGTCTTT 7144  
Db 7075 ACCAGTGGGATGAATCTATATTGTCATCGGATAACACAGGACACCACTGGCTGTGTTC 7134  
Qy 7145 AGTGAACAGGTTTGTATGATCAAGCAAAATCTGTACACAGCCCTATGGGAGATCTACATG 7204  
Db 7135 AGTAGCAATGGGCTTATGCTGAAACAGATTTCAGTACACTGTCATATGGGAAATCTATTTT 7194  
Qy 7205 GATACCAACCCCACTTTTCAGATCATCATAGGCTACCATGGTGGCCCTCTATGATCCACTC 7264  
Db 7195 GACTCTAATATTGACTTTCAACTGGTAATTGGATTTTCATGGTGGCCTGTATGACCCACTC 7254





Db 1678 TCITTTAATACCAATTGTTATAGAGTCTGTGGTGGAAATGTCCCCGAAATTGCCATGGAAAT 1737

QY 1754 GGTGACTGCATCTCTGGGACCTGCCACTGCTTCCTCGGTTCTCTGGGCCCGGCTGTGGC 1813

Db 1738 GGAGAAATGCGTTTCTGGAACTTGCCATTGTTTCCAGGATTTCTGGGTCCGGATTGTTCA 1797

QY 1814 AGAGCCTCTGCCCCGTGCTCTGTAGCGGAAATGGCCAAATACATGAAGGAGAGATGCTTG 1873

Db 1798 AGAGCCGCTGTCCAGTGTTATGTAGTGGCAACGGGCAGTACTCCAAGGGCCGCTGCCTG 1857

QY 1874 TGCCACAGTGGCTGGAAGGCGTGAAGTGCAGTGCAGTGTGCCACCAACCAAGTGTATCGATGTG 1933

Db 1858 TGTTTCAGCGGCTGGAAGGCAACCGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCA 1917

QY 1934 GCCTGAGCAACCATGGCACCTTGCATCACGGGCACCTTGCATCTGCAACCCCTGGCTFACAAG 1993

Db 1918 CAGTGTGGGGTCTGTGGGATTTGTATCATGGGCTCCTGTGCTTGCAGCTCAGGATACAAA 1977

QY 1994 GCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCAACATGTTTCAGGCCGGGTGCTGC 2053

Db 1978 GGAGAAATGTGAAGAAGCTGACTGTATAGACCCCTGGGTGTTCTAATCATGGTGTGT 2037

QY 2054 GTGAGAGCGGAATGCCATTGCTTTGTGGGATGGGAGGCAACCAACTGCGAGACCCCGAGG 2113

Db 2038 ATCCACGGGAATGTCACTGCAGTCCAGGATGGGAGGTAGCAATTGTGAATAACTGAAG 2097

QY 2114 GCCACATGCTTAGACCAGTGTTCAGGCCACGGAACTTCTCCGGACACCCGGSTTTGC 2173

Db 2098 ACCATGTGTCAGACCAGTGTCTCGGCCACGGACGTATCTTCAAGAAAGTGGCTCCTGC 2157

QY 2174 AGCTGTGACCCCAAGCTGGACTGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGT 2233

Db 2158 ACGTGTGACCCCTAACTGGACTGGCCAGACTGCTCAACAGAAATATGTTCTGTGACTGT 2217

QY 2234 GGTGCCATGGCGTGTGCTAGGGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGCA 2293

Db 2218 GGCTCACACGGCGTTTGCAATGGGGGGACGTGTGCTGTGAAGAAGCTGGACGGGCCCA 2277

QY 2294 GCCTGCAACGAGGGGCTGCCACCCGCGCTGTGCCGAGCATGGGACCTGCCCGGACGGC 2353

Db 2278 GCCTGTAAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGACGGGACCTGCAAGGATGGC 2337

QY 2354 AAGTGCAGTGCAGCCCTGGCTGGAATGGCGAAACACTGACCATCGCTCACTATCTGGAT 2413

Db 2338 AAGTGTAAATGCAGCCAGGGCTGGAATGGAGAGCACTGCACTATCGCTCACTATTGGAT 2397

QY 2414 AGGTAGT-----TAAAGAGGGTTGCCCTGGTTGTGCAATGGCAAC 2455

Db 2398 AAGATAGTTAAAGACAAGATAGGATATAAAGAGGGTTGTCTGTGTGCAACAGCAAT 2457

QY 2456 GGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTGTGCCAGCTGGGCTGGAGAGGA 2515

Db 2458 GGAAGATGTACCTTGACCAAAATGGCGGACATTGTGTGCGCAGCCTGGATGGAGAGGA 2517

QY 2516 GCTGGCTGTGACACTTCCATGGAGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCG 2575

Db 2518 GCAGGCTGTGACGTAGCCATGGAGACTCTTTGCACAGATAGCAAGGACAAATGAAGGGAT 2577

QY 2576 GGCCTGGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCG 2635

Db 2578 GGACTCATTGACTGCATGGATCCCGATTGCTGCCCTACAGAGTTCTGCCAGAATCAGCCC 2637

QY 2636 CTGTGCCTTGGCTCCCCTAACCTCTTGACATCATCCAGGAGACACAGGTCCCTGTGTCA 2695

Db 2638 TATTGTGGGGACTGCCGGATCCTCAGGACATCATTAGCCAAAGCCTTCAATCGCCTTCT 2697

QY 2696 CAGCAGAACCTACACTCCTTCTATGACCGCATCAAGTTCCTGTGGGAGGACAGCACG 2755

Db 2698 CAGCAAGCTGCCAAATCCTTTTATGATCGAAATCAGTTTCTTATAGGATCTGATAGCAC 2757

QY 2756 CACATAATCCCGGGAGAACCCCTTTGATGGAGGCGATGCTTGTGTTATTCGTGGCCAA 2815

Db 2758 CATGTTATACCTGGAGAAAGTCCCTTCAATAAGAGCCCTTGCACTGTCTCATCAGAGGCCAA 2817

QY 2816 GTGATGACATCAGATGGAACCCCTCGTGGTGGTGTGAACATCAGTTTTGTCAATAACCCT 2875

Db 2818 GTACTGACTGCTGATGGAACCTCCAATTTGGAGTAAATGTCTCGTTTTCATTAACCA 2877

QY 2876 CTCCTTTGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGG 2935

Db 2878 GAATATGGATATACTATTACCCGCGCAGGACGGAATGTTTGACTTGGTGCAAAATGGTGG 2937

QY 2936 ATCTCCATCATCTCGGTTTCGAGCGGGCACCTTTTCATCACAGGAGCACACCCTGTGG 2995

Db 2938 GCCTCTCTAACTTTGGTATTTGAACGATCCCCATTCCTCACTCAGTATCATACTGTGTGG 2997

QY 2996 CTGCCATGGGATCGCTTCTTTGTGATGGAACCATCATCATGAGACATGAGGAGAAATGAG 3055

Db 2998 ATTCCATGSAATGTCTTTTATGTGATGGATACCTTAGTCTGAGAGAAAGAGAAATGAC 3057

QY 3056 ATTCCCAGCTGTGACCTGAGCAATTTTCCCCGCCCCCAACCCAGTCGTCTCTCCATCCCCA 3115

Db 3058 ATTCCCAGCTGTGATCTGAGTGGATTCGTGAGGCCAAATCCCATCATTTGTGTATCACCT 3117

QY 3116 CTGACGCTCCTTCGCCAGCTCCTGTGCAGAGAAAAGGCCCCCATTTGTCCGGAAAATTCAGGCT 3175

Db 3118 TTATCCACCTTTTTCAGATCTTCTCTGAAGACAGTCCCATCATTTCCGAAAACACAGGTA 3177

QY 3176 TTGCAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGG 3235

Db 3178 CTCCACGAGGAAAACATAATTCAGGAAACAGATTTGAAACTCTCTACTTTGAGTTCAGAA 3237

QY 3236 ACCCTGGCTACAAATCTGTCTGAGGATCAGCCTCACCCACCGACCATCCCTTTCAAC 3295

Db 3238 GCTGACGGTATAAGTCAGTTCTCAAGATCACCATGACCCAGTCTATTTCCATTTAAT 3297

QY 3296 CTCATGAAGTGCACCTCATGGTAGCGGTGGAGGGCCGCTCTTCAGGAAAGTGGTTGCT 3355

Db 3298 TTAATGAAGTTCATCTTATGGTAGCTGTAGTAGGAAGACTCTTCCAAAAGTGGTTTCCT 3357

QY 3356 GCAGCCCCAGACCTGTCTTATTTTCAATTTGGGACAAGACAGACGTCTACAACCAAGAA 3415

Db 3358 GCCTCACCAAACTTGGCCTATACCTTTCATATGGGATAAAACAGATGCATATAATCAGAAA 3417

QY 3416 GTGTTTGGGCTTTCAGAAGCCTTTGTTCCTGCTGGTTATGAATATGAATCCTGCCCAGAT 3475

Db 3418 GTCTATGGTCTATCTGAAGCTGTTGTGTGATGATGATGATGATGATGATGATGATGATGATG 3477

QY 3476 CTAATCCTGTGGAAAAAAGAACACACAGTGTGCAGGGCTATGAATTTGACGCTCCAAG 3535

Db 3478 CTGACTCTGTGGAAAGAGGACTGSCATTTCTGCAGGGCTATGAATGGATCGTCCAAC 3537

QY 3536 CTTGGAGGATGGAGCCTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTGCAC 3595

Db 3538 ATGGTGGCTGGACATTAGATAAACATCACGTGCTGGATGTACAGAACGGTATACTGTAC 3597

QY 3596 AAAGGGAATGGGAGAACCAAGTTTGTCTCAGCAGCCTCCTGTCAATTTGGGAGCATCATG 3655

Db 3598 AAGGAAAACGGGAAAACCAAGTTTCTCTCCAGCAGCCTCCAGTCGTGAGTAGCATCATG 3657

QY 3656 GGCAATGGGCGCGGAGAAAGCATCTCTGCCCCAGCTGCAACGGCCTTGCTGACGGCAAC 3715

Db 3658 GGCAATGGGCGAAGGCGCAGCATTTCTCTGCCCCAGTTGCAATGGTCAAGCTGATGGTAAC 3717

QY 3716 AAGCTCCTGGCCCCAGTGGCCCCCTACCTGTGGCTCTGACGGGAGCCTCTATGTGGTGAT 3775

Db 3718 AAGTTACTGGCCCCAGTGGCGCTAGCTTGTGGGATCGATGGCAGTCTGTACGTAGCGCAT 3777

QY 3776 TTCAACTACATTAGAAGGATCTTCCCTCTGGAAATGTCAACCAACATCCTAGAGCTGAG 3835

Db 3778 TTCAACTACGTGGCGGATATTCCCTTCTGGAATGTAAACAAAGTGTCTTAGAACTAAGA 3837

QY 3836 AATAAAGATTTACAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCC 3895

Db 3838 AATAAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCA 3897

QY 3896 ATGAGTGGGCGCTCTTCTCTTCTGACAGCAACAGCCGGGGTCTTTAAATCAAGTCC 3955  
Db 3898 GTACGGGAGATCTGTACGTTTCTTGACACAACACCCCGCAGAAATTTATCGCCCAAGTCA 3957  
QY 3956 ACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAG 4015  
Db 3958 CTACGGGGCAAAAGACTTGACTAAATATGAGAAGTCTGTCGACGGACAGGGGAGCAA 4017  
QY 4016 TGCCTCCCTTTTGATGACACTCGCTGCGGGGATGGTGGGAAGCCACAGAAGCCACACTC 4075  
Db 4018 TGCCTCCGTTTGACGAGCGGAGATGTGGGATGGAGGGAAGGCCGTGGAAGCCACACTC 4077  
QY 4076 ACCAATCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGTTGATGGCACC 4135  
Db 4078 ATGAGTCCCAAGGAATGGCAGTTGATAAGATGGATTAATCTACTTTTGTGATGGAACC 4137  
QY 4136 ATGATCAGACGCATCGATCAGAATGGGATCATCTCCACCCTGTCTCGGCTCTAATGATCTC 4195  
Db 4138 ATGATTAGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTG 4197  
QY 4196 ACATCAGCCCGGCACCTAGCATCAACCCAAATGGACAACCTCACTTTATGTCTCGACAACAAT 4255  
Db 4198 ACTTCAGCCAGACCTTTAACTTGTGACACCAGCATGCACATCAGCCAGGTACGTCTGGAA 4257  
QY 4256 TGGCCACACAGACTTAGCATCAACCCAAATGGACAACCTCACTTTATGTCTCGACAACAAT 4315  
Db 4258 TGGCCCACTGACCTAGCCATTAAACCCTATGGATAACTCCATTTATGTCTCGGATAATAAT 4317  
QY 4316 GTGGTCTCTGAAAATCTGAAAACCAACAGGTGGGCATTTGTCGCCGGGAGGCCCATGCAAC 4375  
Db 4318 GTAGTTTACAGATCACTGAAAATCGTCAAGTTCGCATTGCTGTCGACGGGCCCATGCAAC 4377  
QY 4376 TGCCAGGTCCCTGGCATTGACCACCTTCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTG 4435  
Db 4378 TGTCAGGTTCCCGGAGTGG--AATATCCTGTGGGAAGCAACGCGGTGCAGACAACACTG 4434  
QY 4436 GAGTCAGCCACCGCTTTGGCTGTTTACACAATGGGGTCCCTGTATATTGCTGAGACTGAT 4495  
Db 4435 GAATCAGCCACTGCCATTGCTGTCTCTACAGTGGGTCCCTGTACATTACTGAAACTGAT 4494  
QY 4496 GAGAAAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCT 4555  
Db 4495 GAGAAGAAATTAACCGGATAAGGCAGGTCAACACAGATGGAGAAATCTCCTTAGTGGCC 4554  
QY 4556 GGGGCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTTCTTGGAGAC 4615  
Db 4555 GGAATACCTTCAGAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTTACCAGAGTGA 4614  
QY 4616 GATGTTATGCCAAGGATGCAAGTTAAATACCCCATCTTCTTGGCTGTGTGCTGAT 4675  
Db 4615 GATGGCTACGCCAAGGATGCCAACTCAGTGGCCCCATCTCTCCCTGGCTGCTTCTCCAGAT 4674  
QY 4676 GGGGAGCTCTACGTGGCGGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAG 4735  
Db 4675 GGTACACTGTATATTCAGATCTAGGGAATATCCGGATCCGGGCTGTGTCAAAGAATAAG 4734  
QY 4736 CCTTCTCAACACCCAGAACATGTATGAGTGTCTTCAACCAATTGACCAAGGAGCTCTAT 4795  
Db 4735 CCTTTACTTAACTCTATGAACTTCTATGAAGTTCGTAAAGTTTAGTCACTGGTGATTACCTT 4794  
QY 4796 CTGTTTGATACCAACCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTG 4855  
Db 4795 ATCTTTGACATCAATGGTACTCACCAATATACTGTAAAGTTTAGTCACTGGTGATTACCTT 4854  
QY 4856 TACAACTTCACTACTGGGACGGCGACATCACACTCATCACAGACAACAATGGCAAC 4915  
Db 4855 TACAATTTTAGCTACAGCAATGACAATGATATTAATGCTGTGACAGACAGCAATGGCAAC 4914  
QY 4916 ATGGTAAATGTCGCGGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGC 4975  
Db 4915 ACCCTTAGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCCTGATAAC 4974  
QY 4976 CAGGTGTACTGGGTGACCATTGGGCACCAACAGTGCACCTCAAGAGTGTGACCCACACAAGGA 5035

Db 4975 CAAGTGATATGGTTGACAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGA 5034  
QY 5036 CACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTTGCAACCAAAAGCAAT 5095  
Db 5035 CTGGAATTAGTTTGTCTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAAGTGAT 5094  
QY 5096 GAAAAACGATGGACAAACATTTTATGAGTACGACAGCTTTGGCCGCTGACAAAATGTGACC 5155  
Db 5095 GAAACTGGATGGACAAACGTTTGTGACTATGACAGTGAAGTCTGCTGACAAAATGTTACG 5154  
QY 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAAGTGATACAGACAGTTTCAGTGCATGTCCAG 5215  
Db 5155 TTTCCAACTGGAGTGTGACAAACCTTGATGGGACATGGACAAGGCTATCACAGTGGAC 5214  
QY 5216 GTAGAGACCTCCAGC--AAGGATGATGTCAACATAACCAACCAACCTGTCTGCCTCAGGC 5272  
Db 5215 ATTGAGTCACTAGCCGAGAAGAAGATGTGAGCATCACTTCAAAATCTGTCTCGATCGAT 5274  
QY 5273 GCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGTACTACATCGGGGCCGATGGC 5332  
Db 5275 TCTTCTACACCATGTTCAAGATCAGTTAAGAAACAGTACCAGATTGGTTATGACGGC 5334  
QY 5333 TCCTTGGCGTGTCTGTGGCCAAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTG 5392  
Db 5335 TCCCTCAGAAATTATCTACGCCAGTGGCCTGGACTCACTACCAAAACAGAGCCGACGTT 5394  
QY 5393 CTGGCTGGCACCGTCAACCCCAACCGTGGGCAAGAGGAATGTCAACGCTGCCCATCGACAAC 5452  
Db 5395 CTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAACATGACTTGCCTGGCGAGAAC 5454  
QY 5453 GGCTCAACCTGTGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTC 5512  
Db 5455 GGTCAAAACTTGTGGAATGGAGATTCCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTC 5514  
QY 5513 TTTGGGCGCGCTGCGGGTGCACAAACGGAATCTCCTATCTCTGGACTTTGATCGCGTA 5572  
Db 5515 TTTGGCCGCAAGCTCAGGGTTAATGGCAGAAAYCTCCTTTAGTTGACTTTGATCGAACA 5574  
QY 5573 ACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTTCAACCTTCGGATTTCTGACGAC 5632  
Db 5575 ACAAAGACAGAAAARATCTATGACGACCAACCGTAAATTTCTACTGAGGATCGCCTACGAC 5634  
QY 5633 CAGCGGGGCGGCCAGCCTCTGTGTACCCAGCAGCAGGCTGAATGGTGTCAACGTGACA 5692  
Db 5635 ACGTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGGCCGTCAATGTCAAC 5694  
QY 5693 TACTCCCTGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTGTGAAAAGAAATGGAA 5752  
Db 5695 TATTATCCACAGGTCAAATTCCAGCATCCAGCGAGGCCACCACTAGCGAGAGAAAGTAGAT 5754  
QY 5753 TACGACCAAGCGGGCCGCATCACATCCAGSATCTTCGCTGATGGGAAGACATGGAGCTAC 5812  
Db 5755 TATGACGGACAGGGGAGGATCGTGTCTCGGTCTTTGCTGATGTTAAACATGGAGTTAC 5814  
QY 5813 ACATACTTAGAAGTCCATGGTGTCTGCTACTACACAGCCAGAGGAGTATATCTTTGAG 5872  
Db 5815 ACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGCCAGCGGAGTACATCTTCGAA 5874  
QY 5873 TTCGACAAGAAATGACCGCCTCTCTTCTGTGACGATGCCCAACGTTGGCGGCGGACACTA 5932  
Db 5875 TACGATATGTGGACCGCCTGTCTGCCATCACCATGCCCCAGTGTGGCTGCCACACCATG 5934  
QY 5933 GAGACCATCCGCTCAGTGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCC 5992  
Db 5935 CAGACCATCCGATCCATTGGTACTACCGCAACATATACAACCCCCCGGAAAGCAACGCC 5994  
QY 5993 TCAGTCATACAGGACTTCACTGAGGATGGGCACCTCTTTCACACCTTCTACCTGGGCAC 6052  
Db 5995 TCCATCATCAGGACTACAACGAGGAAGGCTGCTTCTACAAACAGCTTTCTTGGGTACA 6054  
QY 6053 GGCCGCGAGGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGAC 6112

Db 6055 AGTCGGAGGGTCTTATTCAAATAACAGAAGGCAGACTAGGCTCTCAGAAAATTTTATATGAT 6114

Qy 6113 ACCACCAAGGTCAAGTTTCACTCTATGACGAGACGGCAGGCATGCTGAAGACCATCAACCTA 6172

Db 6115 AGCACAAGAGTCAGTTTACCTATGATGAACACAGCAGGAGTCCTAAAGACAGTAACACCTC 6174

Qy 6173 CAGAATGAGGGCTTCACCTGACACCATCCGCTACCGTCAGATTGGGCCCTGATTGACCGA 6232

Db 6175 CAGAGTGATGGTTTATTGACCATTAGATACAGGCAAAATGGTCCCTGATTGACAGG 6234

Qy 6233 CAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTGGACTACAACTATGAC 6292

Db 6235 CAGATTTCCGCTTTAGTGAAGATGGATGGTAAATGCAAGATTGACTATAGCTATGAC 6294

Qy 6293 AACAGCTTCCGGGTGACCAAGCATGCAGGCTGTGATCAACGAGACCCCACTGCCATTGAT 6352

Db 6295 AACAGCTTTCGAGTGACCAAGCATGCAGGGTGTGATCAATGAACCGCCACTGCCATTGAT 6354

Qy 6353 CTCATCGCTATGATGATGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCATT 6412

Db 6355 CTGATCAGTTTGATGACATTTCTGGCAAGTTTGAGCAGTTTGGAAAGTTTGGAGTTATA 6414

Qy 6413 TACTATGACATTAAACAGATCATCAACAGCTGTGATCAACCAACCAACAGCATTTTGAT 6472

Db 6415 TATTATGATATTAAACAGATCATTTCTACAGCTGTAAATGACCTATACGAAGCAGCTTTGAT 6474

Qy 6473 GCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATG 6532

Db 6475 GCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTCAGGTCGCTCATGTACTGGATT 6534

Qy 6533 ACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTAC 6592

Db 6535 ACAATTCAAGTATGATAACATGGGTCCGGTAAACCAAGAGAGAGATTAAAAATAGGGCCCTTT 6594

Qy 6593 GCCAATACCAGTCACTCGCTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCC 6652

Db 6595 GCCAACACCAACCAATATGCTTATGAATATGATGTTGATGGACAGCTCCAAACAGTTTAC 6654

Qy 6653 ATCAATGACAAGCCACTCTGGGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTG 6712

Db 6655 CTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTGAATGGAAACCTCCATTACTG 6714

Qy 6713 AGCCTGGGAACAGTGCACGGCTCACACCACCTACGGTATGACATCCGCGACCGCATCACT 6772

Db 6715 AACCAAGTAACAGTCGCGCTGACACCCCTTCGCTATGACCTGCGAGACAGAACTCACT 6774

Qy 6773 CGSTGGGTGACGTGCAATACAAGATGGAATGAGGATGGCTTCTTGAGGCGAGGGGCGGT 6832

Db 6775 CGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGTTTCTTACGTCAAGGGGCACG 6834

Qy 6833 GATATCTTTGAGTACAACCTCAGCTGGCCTGCTCATCAAGGCCTACAACCCGGGTGGCAGC 6892

Db 6835 GAAATCTTTGAATATAGCTCAAGGGGCTTCTAACTCGAGTTTACAGTAAAGCGAGTGGC 6894

Qy 6893 TGGAGTGTCAAGTACCGCTACGATGGCCTGGGGCGGCGGTGTCAGCAAGAGCAGCCAC 6952

Db 6895 TGGACAGTGATCTACCGTTATGACGGCCTGGGAAGGCGTGTCTTAGCAAAACCAAGTCTA 6954

Qy 6953 AGCCACCACTGCAGTCTTCTATGCAGACCTGACCAACCCCAAGGTCAACCCACCTG 7012

Db 6955 GGACAGCACCTGCAGTTTTTTTATGCTGACTTAACTTATCCCACTAGGATTACTCATGTC 7014

Qy 7013 TACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGAAGGACACCTCTTT 7072

Db 7015 TACAACCACTCGAGTTCAGAAATACCTCCCTGTATTATGATCTCCAAGGACATCTTTT 7074

Qy 7073 GCCATGGAGCTGAGCAGTGGTGATGATTTTACATAGCTTGTGACAACATCGGACCCCT 7132

Db 7075 GCCATGGAAATCAGCAGTGGGATGAATCTATATTGATCGGATAACACAGGSACACCA 7134

Qy 7133 CTTGCTGTCTTTAGTGGAACAGGTTTATGATCAAGCAAAATCCTGTACACAGCCTATGGG 7192

Db 7135 CTGGCTGTGTTCAGTAGCAATGGGCTTATGCTGAAACAGATTTCAGTACACTGCTGCATATGGG 7194

Qy 7193 GAGATCTACATGGATACCAACCCCAACTTTTCAGATCATCATAGGCTACCATGGTGGCCCTC 7252

Db 7195 GAAATCTATTTGACTCTATATATTGACTTTCAACTGGTAATTGGATTTTCATGGTGGCCCTG 7254

Qy 7253 TATGATCCACTCACCAAGCTTGTCCACATGGGCGGCGGAGATATTATGATGTGCTGGCCGGA 7312

Db 7255 TATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAGAGATTATGACATTTTGGCAGGA 7314

Qy 7313 CGCTGACTAGCCCAGACCAAGAGCTGTGGAAGCACTTAGTAGCAGCAACGTCATGCCT 7372

Db 7315 CGGTGGACAACAACCTGACATAGAAAATCTGGA--AAAGAAATGGGAAGGACCCAGCTCCT 7371

Qy 7373 TTTAATCTCTATATGTTCAAAAACAACAACCCCATCAAGCAACTCCCAGGACATCAAGTGC 7432

Db 7372 TTTAACTTGTAATGTTTAGGAATAACAACCCCTGCAAGCAAAAATCCATGACGTGAAAGAT 7431

Qy 7433 TTCATGACAGATGTTAACAGCTGGCTGCTCACTTTGAAACGGTTTGACCAAGCTCTATGGCTCCACA 7492

Db 7432 TACATCACAGATGTTAACAGCTGGCTGGTGACATTTGGTTTCCATCTGCACAAATGCTATT 7491

Qy 7493 CCTGTTATCCAAACCAAGACATGGATGCCATGGAAACCCCTCCTACGAGCTCATCCACACA 7552

Db 7492 CCTGGATTCCTGTTCCCAAAATTTGATTTAAACAGAACCTTCTTACGA-----A 7539

Qy 7553 CAGATGAAAAACGAGGAGTGGGACAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTA 7612

Db 7540 CTTGTGAAGAGTCAGCAGTGGGATGATATACCGCCCATCTTTCGGAGTCCAGCAGCAAGTG 7599

Qy 7613 CAGAAGCAGCTCAAGGCCCTTTGTCACCTTAGAACGGTTTGACCAAGCTCTATGGCTCCACA 7672

Db 7600 GCGCGGAGGCCAAGGCCCTTCCTGTCGCTGGGGAAGATGSCCGAGGTGCAGGTGAGCCGG 7659

Qy 7673 ATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAAGTTTGCATCCAGCGCTCAGTCTTT 7732

Db 7660 GCGCGGCGCGC--GGCGCGAGTCTCTGGCTGTGGTTGCGCACGGTCAAGTCTGCTGATC 7716

Qy 7733 GGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGCCAGAGTACCACAGACATCATCAGTGTG 7792

Db 7717 GGCAAGGGCGTCATGCTGGCCGTCAGCCAGGGCCGCTGAGACCAACGTCGCTCAACATC 7776

Qy 7793 GCCAATGAGGATGGGCGAAGGGTGTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAAC 7852

Db 7777 GCCAACGAGGACTGCATCAAGGTGGCGCGCTGCTCAACAACGCCTTCTACCTGGAGAAC 7836

Qy 7853 CTGCATCTCAACATTTGATGGGTGGATACCCATTACTTTGTGAACCCAGGACCTTCAGAA 7912

Db 7837 CTGCATCTCAACATCGAGGGCAAGGACACGCATCTTCTCAAGAACCAACCGCCCGAG 7896

Qy 7913 GGTGACCTGGCCATCCTGGGCTCAGTGGGGGGCGGCGAACCCCTGGAGAAATGGGTCAAC 7972

Db 7897 AGCGACCTGGGCACGCTGCGGTTGACAGCGGCGCGCAAGGCGCTGGAGAACGGCATCAAC 7956

Qy 7973 GTCACCTGTGCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATC 8032

Db 7957 GTGACGCTGTCGAGTCCACCACGCTGGTGAACGGCAGGACGCGCAGGTTTCGCGGACGTG 8016

Qy 8033 CAGCTCCAGTACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTTGGATGAGGAG 8092

Db 8017 GAGATGCAGTTCGGCGCGCTGGCGCTGCAGCTGCGCTACGCGATGACCTTGACGAGGAG 8076

Qy 8093 AAGGCACGGGTCTTGAGCTGGCCCGGCAGAGAGCCGTGCGCAAGCGTGGGCCCGCGAG 8152

Db 8077 AAGGCAGCATCTTGAGCAGGCGCGGCAGCGCGCTCGCCCGGCTGGCGCGCGAG 8136

Qy 8153 CAGCAGACACTCGGGGAAGGGAGGAAGGCTCGGGGCTGACAGAGGGGGAGAGCAG 8212

Db 8137 CAGCAGCGCTGCGCGACGCGCAGGAGGGCGCGGCTCTGACAGGAGGGCGAGAGCGG 8196

Qy 8213 CAGGTGCTGAGCACAGGGCGGGTGCAGAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAG 8272

Db 8197 CAGCTGCTGAGCGCGGCAAGGTGACGGGCTACGACGGGTACTACGTACTCTCGGTGGAG 8256

QY 8273 CAGTACCCAGAACTGTCAGACAGCGGCCAACACATCCACTTCATGACAGAGCGAGATG 8332  
|||||  
Db 8257 CAGTACCCGAGCTGCGCAGACAGCGGCCAACACATCCAGTTCCTGCGCAGAGCGAGATC 8316  
|||||  
QY 8333 GCGCGGAGGTGAC 8345  
|||||  
Db 8317 GCGAGGAGGTAAC 8329  
|||||  
RESULT 6  
ABQ82345  
ID ABQ82345 standard; cDNA; 8473 BP.  
XX  
AC ABQ82345;  
XX  
DT 17-DEC-2002 (first entry)  
XX  
DE Human NOV15c encoding cDNA SEQ ID NO:39.  
XX  
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;  
cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;  
antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;  
antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;  
cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;  
stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;  
Pick's disease; vesicular transport disease; cystic fibrosis; goitre;  
diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;  
ulcerative colitis; gastric disorder; duodenal disorder; infection;  
autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;  
rheumatoid arthritis; gene; chromosome 4; ss.  
XX

OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 258..8144  
FT /\*tag= a  
FT /product= "NOV15c"  
XX  
XX WO200262999-A2.  
XX  
PD 15-AUG-2002.  
XX  
XX 31-DEC-2001; 2001WO-US049976.  
XX  
XX 29-DEC-2000; 2000US-0258928P.  
PR 02-JAN-2001; 2001US-0259415P.  
PR 04-JAN-2001; 2001US-0259785P.  
PR 20-FEB-2001; 2001US-0259814P.  
PR 09-MAR-2001; 2001US-0279863P.  
PR 29-MAR-2001; 2001US-0279832P.  
PR 29-MAR-2001; 2001US-0279833P.  
PR 13-APR-2001; 2001US-0283889P.  
PR 18-APR-2001; 2001US-0284447P.  
PR 25-APR-2001; 2001US-0286683P.  
PR 29-MAY-2001; 2001US-0294080P.  
PR 16-AUG-2001; 2001US-0312915P.  
PR 17-AUG-2001; 2001US-0313325P.  
PR 17-SEP-2001; 2001US-0322699P.  
PR 26-NOV-2001; 2001US-0333350P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;  
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;  
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;  
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;  
PI Gunther E, Smithson G, Millet I, Macdougall JR;  
XX  
XX WPI; 2002-732706/79.  
DR P-PSDB; ABP53588.  
XX  
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-  
associated disorders, such as cancers, neurological disorders, disorders  
PT

PT of vesicular transport, gastrointestinal disorders, and autoimmune  
PT diseases.  
XX  
PS Claim 8; Page 119-121; 444pp; English.  
XX  
CC The present invention describes novel human proteins designated NOVX,  
where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,  
cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,  
antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,  
immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,  
protozoacide and antihelminthic activities, and can be used in gene  
therapy. The NOVX proteins, nucleotides or antibodies can be used in the  
manufacture of a medicament for treating a syndrome associated with a  
human disease selected from NOVX-associated disorder, such as cancers  
(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,  
ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,  
ischaemic cerebrovascular disease, Alzheimer's disease or Pick's  
disease), disorders of vesicular transport (e.g. cystic fibrosis,  
diabetes mellitus, Grave's disease, or goitre), gastrointestinal  
disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),  
autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic  
anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic  
and protozoal infections. The NOVX proteins can be used as immunogens to  
produce antibodies and as vaccines. The NOVX nucleotide sequences may be  
used in chromosome mapping, identifying individuals from minute  
biological samples (tissue typing), and in forensic identification of a  
biological sample. The present sequence encodes human NOV15c, which is  
located on chromosome 4  
XX  
SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Query Match 41.0%; Score 3424.4; DB 6; Length 8473;  
Best Local Similarity 67.1%; Pred. No. 0;  
Matches 5027; Conservative 1; Mismatches 2369; Indels 93; Gaps 9;  
QY 880 CGGGCACTTCCTCTCAAGCCTGG---AGGCACCTCCCGCTCTTCTGCACCATCACC 936  
|||  
Db 725 CAGGCATTTCTTATTCAAACACAGGAACAGGTACACGCCACTGTTCACTACTGCAACCCC 784  
|||  
QY 937 AGGTACCCCACTGACGTCCAGCACAGTGTACTCTCTCGCGCCCGACCCCTGCCCGCAG 996  
|||  
Db 785 AGGATACACAATGGCATCTGGCTCTGTTATTACCCACCTACTCGGCCACTACCTAGAAA 844  
|||  
QY 997 CACCTTCGCGCGCGCCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAAGTGAAGTG 1056  
|||  
Db 845 CACCCCTATCAAGAAGTGCTTTTAAATTCAGAAGTCTTCAAGTACTGTAGTGGAAATG 904  
|||  
QY 1057 GCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCTATCTGCTGGCATACTTTGT 1116  
|||  
Db 905 CACTGCACTGTGCGCGTAGGGGTCTCGGTCTCTCGGCAATACTCTCTTATTTAT 964  
|||  
QY 1117 GGCCATGCACCTGTTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTATGA 1176  
|||  
Db 965 AGCAATGCATCTCTTTGGCCTCAACTGGCAGCTACAGCAGACTGAAAATGACACATTTGA 1024  
|||  
QY 1177 GATCAGGAGGACACAGCCAGAGTTGGCCTGTGCCAACCGAGCTCTCCCTATACCCCTC 1236  
|||  
Db 1025 GA-----ATGGAAGTGAATTTGATATCCATGCCAACAAACACTGTGTCTATTACCTTC 1078  
|||  
QY 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGGAAGGCAAGAACCAAGAAAGGCC 1296  
|||  
Db 1079 TGGAGACA-----ATGGAAGT 1096  
|||  
QY 1297 CAGTAGTTTCTTCCAGAGGACAGTTTCATAGATTCTGGAGAAATGTATGTGGGAGCGG 1356  
|||  
Db 1097 AGGTGGATTTCGCAAGAAAATAACACCATAGATTCCGGAGAACTTGATATTGGCCGAAG 1156  
|||  
QY 1357 AGCCTCCCAAGAGATTCTCTCGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416  
|||  
Db 1157 AGCAATTCAGAGATTCTCTCCCGGATCTTCTGGAGATCACAGCTCTTCATTGATCAGCC 1216  
|||  
QY 1417 TGTGCATCTGAAATTCATGTCTCTGGGAAAGGACGCCCTGGTGGCATTATGGCAG 1476  
|||

Db 1217 ACAGTTTCTTAAATTCAATATCTCTCTTCAGAAAGGATGCATTGATTGGAGTATATGGCCG 1276

Qy 1477 AAAAGSCTCCCTCCTTCACATACACAGTTTGTGACTTTGTGGAGCTGCTGGATGGCAGGAG 1536

Db 1277 GAAAGGTTACCGCTTCCCATACTAGTATGACTTCTGGAGCTCCTGGATGGCAGCAG 1336

Qy 1537 GCTCCTAAACCCAGGAGGCGGAGCCCTAGAGGGGACCCCGCCAGTCTCGGGAACGTGT 1596

Db 1337 GCTGATTGCCAGAGAGCAGCGGAGCCTGCTTGAGACGAGAGAGCCGGCGGAGGCGGAG 1396

Qy 1597 GCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTGGATTTCAGGAATTCGGCACTT 1656

Db 1397 ATCCGTAGCCTTCATGAGGCGGCTTTATCCAGTACTTGGATTCTGGAATTCGGCATCT 1456

Qy 1657 GCGTTTTCAAATGACGGAAGGAGTCAGAAAGTGGTTTCTTCTCACCACTGCCATTGA 1716

Db 1457 GCGTTTTCATGATGGGAAAATGCAGAGCAGGTGTCTTTTAATACCAATTGTATAGA 1516

Qy 1717 GTCGTGGATAAAGTCCCGCCAGCAACTGCTATGGCAATGGTGAAGTGCATCTCTGGGACCTG 1776

Db 1517 GTCGTGGTGAATGTCCCGGAAATGCCATGGAAATGGAGAATGCGTTTCTGGAACTTG 1576

Qy 1777 CCAGTGTCTTCTGGGTTTCTGGGCCCCGACTGTGGCAGAGCCTCCTGCCCGCTGCTCTG 1836

Db 1577 CCATTGTTTCCAGGATTTCTGGTCCGGAATGTTTCAAGAGCCGCTGTCCAGTGTATG 1636

Qy 1837 TAGCGGAAATGGCCAAATACATGAAGGCGAGATGCTTGTGCCACAGTGGCTGGAAAGCGC 1896

Db 1637 TAGTGGCAACGGGCAGTACTCCAAGGGCCGCTGCCGTGTGTTTTCAGCGGCTGGAAAGGCAC 1696

Qy 1897 TGAGTGGCATGTGCCCAACCAACAGTGTATCGATGTGGCTGCAGCAACCATGGGACCTG 1956

Db 1697 CGAGTGTGATGTCCGACTACCCAGTGTATTGACCCACAGTGTGGGGTCTGGGATTG 1756

Qy 1957 CATCACGGCACCTGCATCTGCAACCTTGGCTACAAGGCGAGAGCTGTGAGGAAGTGA 2016

Db 1757 TATCATGGGCTCTTGTGCTTGCRAACTCAGGATACAAGGAAAAAGTTGTGAAGAAGCTGA 1816

Qy 2017 CTGCATGACCCACATGTTTCAGGCGGGGTGTCTGCGTGAGAGGCGAATGCCATTGCTT 2076

Db 1817 CTGTATAGACCTTGGGTGTTCTTAATCATGGTGTGTGTATCCACGGGGAATGTCACTGCAG 1876

Qy 2077 TGTGGATGGGAGGCACCAACTGCGAGACCCCGAGCTGTGAGCCAGTGTTC 2136

Db 1877 TCCAGGATGGGAGGTAGCAATTGTGAATACTGAAGACCATGTGTCCAGACCAAGTCTC 1936

Qy 2137 AGCCACGGAACTTCTCCTCCGGACACCGGGCTTTGACGTGTGACCCCAAGCTGGAAGTGG 2196

Db 1937 CGGCCACGGAACGTATCTTCAAGAAAGTGGCTCCTGCAAGTGTGACCCCTAACTGGACTGG 1996

Qy 2197 ACAGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGTGGCCCATGGCGTGTGCTAGG 2256

Db 1997 CCCAGACTGCTCAACGAAATATGTTCTGTGGACTGTGGCTCACACGGCGTTTGCATGGG 2056

Qy 2257 GGGCACTGCGCGTGCAGAGGATGGCTGGATGGGGCAGCTGCGACCCAGCGGGCCTGCCA 2316

Db 2057 GGGGACGTGTGCTGTGAAGAAGGCTGGACGGGCCCGCCAGCTGTAATCAGAGAGCCTGCCA 2116

Qy 2317 CCCGCGTGTGCGGAGCATGGGACCTGCCGACGCGCAAGTGGAGTGCAGCCCTGGCTG 2376

Db 2117 CCCCCGTGTGCGGAGCACGCGACCTGCAAGGATGGCAAGTGTGAATGCAGCCAGGCTG 2176

Qy 2377 GAATGGCAACACTGCACCATCGCTCACTATCTGGATAGGTAAGTAAAGACAAGATAGG 2421

Db 2177 GAATGGAGAGCACTGCACATATCGCTCACTATTTGGATAAGATAGTTAAAGACAAGATAGG 2236

Qy 2422 ---TAAAGAGGTTGCCCTGGTTGTGCAATGGCAACGGCAGATGTACCTTAGACCTGAA 2478

Db 2237 ATATAAAGAGGTTGTCTTGTGCTGTGCAACAGCAATGGAAGATGTACCTTGGACCCAAA 2296

Qy 2479 TGGTGGCACTGCGTCTGCCAGCTGGCTGGAGGAGCTGGCTGTGACACTTCCATGGA 2538

Db 2297 TGGCGGACATTGTGTGTGCCAGCCTGATGGAGAGGAGGAGGCTGTGACGTAGCCATGGA 2356

Qy 2539 GACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCTGCTGGACTGCATGGACCC 2598

Db 2357 GACTCTTTGCACAGATAGCAAGGACAATGAAGGGATGGACTCATTTGACTGCATGGATCC 2416

Qy 2599 TGACTGTGCTGCCCTCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCTAAACCC 2658

Db 2417 CGATTGTGCTACAGAGTTCTTGCCAGAATCAGCCCTATTGTGCGGAGACTGCCGGATCC 2476

Qy 2659 TCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGAGAGAACTACATCCTCTTA 2718

Db 2477 TCAGGACATCATTAGCCAAAGCCTTCAATCGCCTTCTCAGCAAGCTGCCAAATCCTTTTA 2536

Qy 2719 TGACCGCATCAAGTTTCTCGTGGCAGGGACAGCACGCACATAAATCCCGGGAGAAACCC 2778

Db 2537 TGATCGAATCAGTTTCTTATAGGATCTGATAGCACCCCATGTTATACCTGGAGAAAGTCC 2596

Qy 2779 CTTTGTATGGAGGGCATGCTTGTGTATTTCGTGGCCAAAGTGTGATGACATCAGATGGAACCC 2838

Db 2597 TTTTCAATAAGAGCCTTGCAATCTGTGATCAGAGGCCCAAGTACTGACTGCTGATGGAACCTC 2656

Qy 2839 CCTGTTGTTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATATACAATCAGCAG 2898

Db 2657 ACTTATGGAGTAAATGTCTCGTTTTCCATTACCCAGAAATGATGATATACTATTACCCG 2716

Qy 2899 GCAAGATGCGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCTCGGTTTGA 2958

Db 2717 NCAGGACGGAATGTTTGACTTGGTGGCAAAATGGTGGGCCCTCNCCTAACTTTGGTATTGA 2776

Qy 2959 GCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTTCTTTGT 3018

Db 2777 ACGATCCCCCATCTCCTCACTCAGTATCATACTGTGTGGATTCATGGAATGTCTTTTATGT 2836

Qy 3019 CATGGAAACCATCATCATGAGACATGAGGAGAAATGAGATTTCCAGCTGTGACCTGAGCAA 3078

Db 2837 GATGGATACCCCTAGTCAATGGAGAAAGAGAAATGACATTTCCAGCTGTGATCTGAGTGG 2896

Qy 3079 TTTTGGCCCCCACAACCCAGTCTGTCTCTCCATCCCCACTGACGTCTTTCGCCAGCTCCTG 3138

Db 2897 ATTCGTGAGGCCAAATCCCATCATTTGTGTCTCATCACCTTTATCCACCTTTTTCAGATCTTC 2956

Qy 3139 TGCAGAGAAAGGCCCCCATTTGTGCCGGAATTTGAGGCTTTTGAGAGGAAATCTCTATCTC 3198

Db 2957 TCCCTGAAGACAGTCCCATCATTTCCCGAAAACACAGGTACTCCACAGGAAACTACAATTCC 3016

Qy 3199 TGGTGCAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCCTGGCTACAAATCTGTCTC 3258

Db 3017 AGGAACAGATTTGAAACTCTCCTACTTGAGTTCAGAGCTGCAGGCTGAGGCTATAAGTCAGTTCT 3076

Qy 3259 GAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGTT 3318

Db 3077 CAAGATCACCATGACCCAGTCTATTATTCCATTAAATTAATGAAGGTTTCATCTTATGTT 3136

Qy 3319 AGCGGTGGAGGCGCCCTCTTTCAGGAAAGTGGTTCGCTGCAGCCCGACACCTGTCTCTATTA 3378

Db 3137 AGCTGTAGTAGGAAGACTCTTCCAAAAGTGGTTTCTCTGCCCTCACCAAACTTGGCCTATAC 3196

Qy 3379 TTTTCATTTGGACAAGACAGACGTCTACAAACCAGAGGTGTTTGGGCTTTTCAGAAGCCTT 3438

Db 3197 TTTTCATATGGGATAAACAGATGCATATAATCAGAAAGTCTATGTTCTATCTGAAGCTGT 3256

Qy 3439 TGTTCCTGGTGGTTATGAATATGAATCCTGTCCCCAGATCTAATCCTGTGGGAAAAAGAAC 3498

Db 3257 TGTGTCAAGTTGGATATGAGTATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAAAGAGGAC 3316

Qy 3499 AACAGTGTGCAGGGCTATGAAATTTGACCGCTCCAAGCTTGGAGGATGGAGCCCTAGACAA 3558

Db 3317 TGCCATTTGACAGGGCTATGAATTTGGATGCGTCCAACATGGGTGGCTGGACATTAGATAA 3376

Qy 3559 ACATCATGCCCTCAACATTTCAAAGTGGTATCTCTGCACAAAGGGAATGGGGAGAACAGTT 3618

Db 3377 ACATCACGTGCTGGATGTACAGAACGGTATACTGTACAAGGGAACGGGGAAAAACAGTT 3436

QY 3619 TGTGCTCAGCAGCCTCCTGTCATTTGGGAGCATCATGGGCAATGGGCGCGGAGAGCAT 3678  
Db 3437 CATCTCCAGCAGCCTCCAGTCGTGAGTAGCATCATGGGCAATGGGCGAAGCGCAGCAT 3496  
QY 3679 CTCTGCCCCAGCTGCAACCGCCTTGCTGACGGGCAACAAGCTCCTGGCCCCAGTGGCCCT 3738  
Db 3497 TTCTGCCCCAGTTGCAATGGTCAAGCTGATGGTAACAAGTTACTGGCCCCAGTGGCGCT 3556  
QY 3739 CACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTT 3798  
Db 3557 AGCTTGTGGGATCGATGGCAGTCTGTACGTAGGCGATTTCAACTACGTCGGCGGATATT 3616  
QY 3799 CCCCCTCTGGAAATGTCACCAACATCCTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCA 3858  
Db 3617 CCTTCTGGAAATGTAACAAGTGTCTTAGAACTAAGAAATAAAGATTTTAGACATAGCAG 3676  
QY 3859 CAGTCCAGCACACAAATATACCTGGCCACAGACCCCATGAGTGGGGCGCTCTTCCCTTC 3918  
Db 3677 CAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACGGGAGATCTGTACGTTTC 3736  
QY 3919 TGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGTGGTGTGAAGGACCTTGT 3978  
Db 3737 TGACACAAACACCCGCAAGATTTATCGCCCAAGTCACTTACGGGGCAAAAGACTTGAC 3796  
QY 3979 CAAGAACTCTGAGGTGTTGCGGGGACAGGTGACAGTGCCTCCCTTTGATGACACTCG 4038  
Db 3797 TAAAAATGCAGAAAGTCGTCGAGGACAGGGGAGCAATGCCTTCCGTTTGACGAGGCGAG 3856  
QY 4039 CTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCAGGGGTATTACAGT 4098  
Db 3857 ATGTGGGATGGAGGAAGCCCGTGGAAAGCCACACTCATGATCCCAAGGAATGGCAGT 3916  
QY 4099 GGACAAGTTTGGGCTGATCTACTTCGTGGATGCGACCATGATCAGACGCATCGATCAGAA 4158  
Db 3917 TGATAAGAATGGATTAACTACTTTGTTGATGGAACCATGATTAGGAAGTTGACCAAAA 3976  
QY 4159 TGGGATCATCTCCACCTCTCGGCTCTAATGATCTCATATCAGCCCGCCACTCAGCTG 4218  
Db 3977 TGGAAATCATCAACTCTCTGGGCTCTAACGATTTGACTTCAGCCAGACCTTTAACTTG 4036  
QY 4219 TGATTCTGTCATGGATATTTCCAGGTAAGACTGGAGTGGCCCAACAGACTTAGCCATCAA 4278  
Db 4037 TGACACCAGCATGCACATCAGCCAGGTACGTCTGGAATGGCCCACTGACCTAGCCATTAA 4096  
QY 4279 CCCAATGGACAACTCACTTTATGTCTCTGCAACAATGTGGTCCCTGCAAAATCTTGAAAA 4338  
Db 4097 CCCTATGGATAACTCCATTTATGTCTCTGGATAATAATGTAGTTTTACAGATCACTGAAA 4156  
QY 4339 CCACCAGGTGCGCATTTGTCGGGAGGCCATGCCATGCATGCCAGGTCCCTGGCATTGACCA 4398  
Db 4157 TCGTCAAGTTGCGATGTCTGCTGGACGGGCCATGCATGTGAGTTCCCGGAGTGG---A 4213  
QY 4399 CTTCTCTGCTAAGCAGGTGGCCATCCACGCAACCTTGGAGTCAAGCCGCTTTGGCTGT 4458  
Db 4214 ATATCTGTGGGGAAGCACGCGGTGCAGAACACACTGGAATCAGCCACTGCCATTGCTGT 4273  
QY 4459 TTCAACAATGGGTCCTGTATATTGCTGACACTGATGAGAAAAAGATCAACCGCATCAG 4518  
Db 4274 GTCTACAGTGGGTCCTGTACATTACTGAAACTGATGAGAAGAAATTAACCGGATAAG 4333  
QY 4519 GCAGGTCAACACTAGTGGAGAGATCTCACTCGTTGCTGGGCCCCAGTGGCTGTGACTG 4578  
Db 4334 GCAGGTCAACAACAGATGGAGAAATCTCCTTAGTGGCCGGAATACCTTCAGAGTGTACTG 4393  
QY 4579 TAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTATGCAAGGATGCAAA 4638  
Db 4394 CAAAAATGATGCCAACTGTGACTGTTACAGAGTGGAGATGGCTACGCCAAGGATGCCAA 4453  
QY 4639 GTTAAATACCCCACTCTCCTTGGCTGTGTGCTGATGGGGAGCTTACGTTGGCCGACCT 4698  
Db 4454 ACTCAGTGGCCCCATCCTCCTGGCTGCTTCCAGATGGTACACTGTATATTGCAGATCT 4513  
QY 4699 TGGGAACATCCGAATTCGGTTTATCCGGGAAGAACAGCCTTTCTCCTCAACCCAGAACAT 4758

Db 4514 AGGGAATATCCGGATCCGGGCTGTGTCAAAAGAAATAAGCCTTTACTTAACCTCTATGAACCTT 4573  
QY 4759 GTATGAGCTGTCTTACCAATTTGACCAGGAGTCTATCTGTTTATACCAACCGGCAAGCA 4818  
Db 4574 CTATGAAGTTGCGTCTCAACTGATCAAGAACTCTACATCTTTGACATCAATGGTACTCA 4633  
QY 4819 CCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACCTTCACTACACTGGGA 4878  
Db 4634 CCAATATACTGTAAGTTTAGTCACTGGTGAATACCTTTACAATTTTAGCTACAGCAATGA 4693  
QY 4879 CGCGGACATCACACTCATCACAGACAAACAATGGCAACATGGTAAATGTCCGCCGAGACTC 4938  
Db 4694 CAATGATATTACTGCTGTGACAGACAGCAATGGCAACACCTTAGAATTAGACGGGACCC 4753  
QY 4939 TACTGGGATGCCCTCTGGCTGGTGTCCAGATGGCCAGGTGTACTGSGGTGACCATGG 4998  
Db 4754 AAATCGCATGCCAGTTCGAGTGGTGTCTCCTGATAACCAAGTGATATGTTGACAATAGG 4813  
QY 4999 CACCAACAGTGCACACTCAAGAGTGTGACCACACAGGACACAGAGTTGGCCATGATGACATA 5058  
Db 4814 AACAAATGGATGTTGAAAGGCATGACTGTCTCAAGGACTGGAAATTTGTTTACTTA 4873  
QY 5059 CCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAAACGGATGGACAACTTTTA 5118  
Db 4874 CCATGGCAATAGTGGCCTTTTAGCCACTAAAAGTGTGAAACTGGATGGACAACTTTT 4933  
QY 5119 TGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCTACTTGGCCAGGTGAGCAG 5178  
Db 4934 TGACTATGACAGTGAAGTCTGTGACAAATGTTACGTTTCCAACCTGAGTGGTTCACAAA 4993  
QY 5179 TTTCCGAAGTGTACAGACAGTTTCAAGTGTCCAGGTAGAGACTCCAGC---AAGGA 5235  
Db 4994 CCTGCATGGGGACATGGACAAGGCTATCACAGTGGACATTGAGTCTATAGCCGAGAAGA 5053  
QY 5236 TGATGTACCCATAACCAACCACTGTCTGCTCAGGCGCTTCTACACACTGCTGCAAGA 5295  
Db 5054 AGATGTGAGCATCACTTCAAATCTGTCTCGATCGATTCGTTTCTTCTACACCATGGTTCAAGA 5113  
QY 5296 CCAAGTCCGGAACAGCTACTACATCGGGGCGGATGGTCTCTTGGGCTGTCTGGCCAA 5355  
Db 5114 TCAGTTAAGAAACAGCTACAGATTGGTTATGACGGCTCCCTCAGAATTTATCTACGCCAG 5173  
QY 5356 CGGCATGGAGGTGGCGCTGCAGACTGAGCCCACTTGTGCTGGTGGCACCCGTCAACCCAC 5415  
Db 5174 TGGCCTGGACTCACACTACCAACAGAGCGCCAGCTTCTGGCTGGCACCGCTAATCCGAC 5233  
QY 5416 CGTGGCAAGAGGAATGTACGCTGCCATCGACAAACGGCTCAACCTGGTGGAGTGGCG 5475  
Db 5234 GGTGCCAAAAGAAAACATGACTTTGCTGCGGAGAACGGTCAAACTTGGTGAATGGAG 5293  
QY 5476 CCAGCGCAAAGAGCAGGCTCGGGGCCAGTCACTGTCTTTGGGCGCGGCTCGGGGTGCA 5535  
Db 5294 ATTCCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAA 5353  
QY 5536 CAACCGAAATCTCTATCTCTGACTTTGATCGCGTAAACGACAGAGAGATCTATGA 5595  
Db 5354 TGGCAGAAACCTCTCTTCACTTGTGATCGAACAAACAGACAGAAAAGATCTATGA 5413  
QY 5596 TGACCACCGCAAGTTCACTTCCGATTTCTGACGACCGAGGCGGGCGCCAGCCTCTG 5655  
Db 5414 CGACCACCGTAAATTTCTACTGAGGATCGCCTACGACAGCTGCGGCACTCTCTG 5473  
QY 5656 GTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGGTACATTGC 5715  
Db 5474 GCTGCCAAGCAGCAAGCTGATGGCCGTCAATGTCACTTTATCCACAGGTCAAATTGC 5533  
QY 5716 TGGCATCCAGAGGGCATCATGTCTGAAAGAAATGGAATACGACAGGCGGGCCGATCAC 5775  
Db 5534 CAGCATCCAGCAGGCAACCACTAGCGAGAAAGTAGATTATGACGACAGGGAGGATCGT 5593  
QY 5776 ATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACATACTTAGAGAAGTCCATGTT 5835

Db 5594 GTCTCGGGTCTTTGCTGATGGTAAAAACATGGAGTTACACATATTTAGAAAAAGTCCATGGT 5653

QY 5836 GCTGCTACTACACAGCCAGAGGCAGTATATCTTTGAGTTCGACAAGAAATGACCGCCTCTC 5895

Db 5654 TCTTCTGCTTCATAGCCAGCGGCAGTACATCTTCGAATACGATATGTGGGACCGCCTGTC 5713

QY 5896 TTCTGTGACGATGCCCAACGCTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGCTA 5955

Db 5714 TGGCATCAACCATGCCCAGTGGCTCGCCACACCATCGAGACCATCCGATCCATGGCTA 5773

QY 5956 CTACAGAAACATCTATCAGCCCCCTGAGGGCAATCGCTCAGTCAATACAGGACTTCACTGA 6015

Db 5774 CTACCGCAACATATACAACCCCCCGGAAAGCAACGCCCTCCATCATCACGGACTACAACGA 5833

QY 6016 GGATGGGCACCTCCTTCACACCTTCTACCTGGGCACCTGGCCGAGGTGATATACAAGTA 6075

Db 5834 GGAAGGCTGCTTCTACAAACAGCTTCTTGGGTACAAGTCGGAGGCTTTATTCAAATA 5893

QY 6076 TGGCAAACTGTCAAAGCTGSCAGAGACGCTCTATGACACCAACCAAGGTCAGTTTCACCTA 6135

Db 5894 CAGAAGGCAGACTAGGCTCTCAGAAATTTTATATGATAGCACAAAGTTCAGTTTACCTA 5953

QY 6136 TGACGAGACGGCAGGCATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACCTGCAC 6195

Db 5954 TGATGAAACAGCAGGAGTCTCTAAAGACAGTAAACCTCCAGAGTGATGGTTTTATTGCAC 6013

QY 6196 CATCCGCTACCGTCAGATTGGGCCCCCTGATTGACCGACAGATCTTCGCTTCACTGAGGA 6255

Db 6014 CATTAGATACAGGCAAAATGGTCCCCTGATTGACAGGCAGATTTTCGCTTTAGTGAAGA 6073

QY 6256 AGGCAATGCTCAACGCCCGTTTGTACTACAACATAAGACACAGCTTCGCGGTGACCAAGCAT 6315

Db 6074 TGGGATGGTAAATGCAAGATTTGACTATAGCTATGACAAACAGCTTTTCGAGTGACCAAGCAT 6133

QY 6316 GCAGGTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTC 6375

Db 6134 GCAGGTGTGATCAATGAAACGCCACTGCCCTATTGATCTCTATCAGTTTGTATGACATTTT 6193

QY 6376 AGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTTAACCAAGATCAT 6435

Db 6194 TGGCAAAAGTTGAGCAGTTTGGAAAGTTTGGAGTTATATATATGATATTAACCAAGATCAT 6253

QY 6436 CACCACAGCTGTCAATGACCCACACCAAGCATTTTGTATGCATATGGCAGGATGAAGGAAGT 6495

Db 6254 TTCTACAGCTGTAATGACCTATACGAAGCACTTTGTATGCTCATGGCCGTATCAAGGAGAT 6313

QY 6496 GCAGTATGAGATCTTCCGCTCGCTCATGTACTGTGATGACCGTCCAGTATGATAACATGGG 6555

Db 6314 TCAATATGAGATATTAGGTGCTCATGTACTGGATTACAATTCAGTATGATAACATGGG 6373

QY 6556 GCGAGTAGTGAAGAAGGAGCTGAAGTAGGACCCCTACGCCAATACCACCTCGCTACTCCTA 6615

Db 6374 TCGGGTAACCAAGAGAGAGATTAAATAAGGCCCTTTTGCCACACACCACCAATATGCTTA 6433

QY 6616 TGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCBAATGACAAGCCACTCTGGCG 6675

Db 6434 TGARTATGATGTTGATGGACAGCTCCAAAACAGTTTACCTCAATGARAAGATAATGTGGCG 6493

QY 6676 CTACAGCTACGACCTCAATGGGAACCTGCACATTACTGAGCCCTGGGAACAGTGCACCGCT 6735

Db 6494 GTACAACTACGATCTGAATGGAAACCTCCATTTACTGAACCCAAAGTAACAGTGGCGCTCT 6553

QY 6736 CACACCACTACGGTATGACATCCGGACCGCATCACTCGGCTGGGTGACGTGCAATACAA 6795

Db 6554 GACACCCCTTCGCTATGACCTGCGAGACAGAATCACTCGACTGGGTGATGTTTCAATATCG 6613

QY 6796 GATGGATGAGGATGGCTTCTTGGGACGGGGCGGTGATATCTTTTGTAGTACAACACTCAGC 6855

Db 6614 GTTGGATGAGATGGTTTCTCTACGTCAAAGGGGCACGGAAATCTTTGAATATATAGCTCCAA 6673

QY 6856 TGGCCTGCTCATCAAGGCCCTACAACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGA 6915

Db 6674 GGGGCTTCTAACTCGAGTTTACAGTAAAGGCAGTGGCTGGACAGTGTACTACCGTTATGA 6733

QY 6916 TGGCCTGGGGCGGCGGTGTCCAGCAAGAGCAGCCACAGCCACCACTGCAGTTCCTCTA 6975

Db 6734 CGGCCGTGGGAAGGCGTGTCTTAGCAAAACCAGTCTAGGACAGCACCTGCAGTTTTTTTA 6793

QY 6976 TGCAGACCTGACCAACCCACCAAGGTCAACCCACCTGTACAACCACTCCAGCTCTGAGAT 7035

Db 6794 TGCTGACTTAACCTTATCCCACTAGGATTACTCATGTGTACAACCATTCGAGTTCAGAAAT 6853

QY 7036 CACCTCCCTCTACTACGACTTGAAGGACACCTCTTTTGCATGGAGCTGAGCAGTGGTGA 7095

Db 6854 TACCTCCCTGTATTATGATCTCAAGGACATCTTTTTTGCATGGAAATCAGCAGTGGGA 6913

QY 7096 TGAGTTTACATAGCTTGTGACAAACATCGGGACCCCTCTTTGTGTCTTTTAGTGAACAGG 7155

Db 6914 TGAATTTCTATATTGCATCGGATAACACAGGGACACCACTGGCTGTGTTTTCAGTAGCAATGG 6973

QY 7156 TTTGTATGATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTACATGGATACCAACCC 7215

Db 6974 GCTTATGCTGAAACAGATTCAGTACACTGCATATGGGGAATCTATTTTGACTCTAATAT 7033

QY 7216 CAACCTTCAGATCATATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGT 7275

Db 7034 TGACTTTCAACTGGTAATTGGATTTCATGGTGGCCTGTATGACCCACTCACCAAAATTAAT 7093

QY 7276 CCACATGGGCGGCGAGATTATGATGTGCTGGCCGGACGCTGGACTAGCCCAGACCCACGA 7335

Db 7094 CCACCTTTGGAGAAAAGAGATTATGACATTTTGGCAGGACGGTGGACAACACCTGCACATGA 7153

QY 7336 GCTGTGGAAGCACCTTAGTAGCAACAGTCAATGCCCTTTTAAATCTCTATATGTTTCAAAAA 7395

Db 7154 AATCTGGA--AAAGAATTGGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAA 7210

QY 7396 CAACAACCCCATCAGCAACTCCCAGGACATCAAGTGTCTTCAAGACAGATGTTAAACAGCTG 7455

Db 7211 TAAACAACCTGCAAGCAAAATCCATGACGTGAAAGATTACATCACAGATGTTAAACAGCTG 7270

QY 7456 GCTGCTCACCTTTGGATTCCAGCTACACAACAGTGTATCCCTGGTTATCCCAAAACAGACAT 7515

Db 7271 GCTGGTGACATTTGGTTTCCATCTGCACAATGCTATTCTCTGGATTCCCTGTTCCCAAAAT 7330

QY 7516 GGATGCCATGGAACCCCTCCTACGAGCTCATCCACACACAGATGAAAACGCGAGGTGGGA 7575

Db 7331 TGATTTAACAGAACCTTCTTACGA-----ACTTGTGAAGAGTCAGCAGTGGGA 7378

QY 7576 CAACAGCAAGTCTATCTCTCGGGGTACAGTGTGAAGTACAGAAGCAGACTCAAGGCCTTTGT 7635

Db 7379 TGATATACCGCCCATCTTCGGAGTCCAGCAGCAAGTGGCGGCGCAGGCCAAGGCCTTCCT 7438

QY 7636 CACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACAATCACCAAGTGTCCAGCAGGCTCC 7695

Db 7439 GTCGTGGGGAAGATGGCCGAGGTGCAAGTGAGCCGGCGCGGGCGGGC---GGCGCGCA 7495

QY 7696 AAAGACCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTT 7755

Db 7496 GTCCTGGCTGTGTTTCGCCACGGTCAAGTCGCTGATCGGCAAGGGCGTCAATGCTGGCCGT 7555

QY 7756 GAAGGATGGCCAGTGACCACAGACATCATCAGTGTGGCCAATGAGATGGGCGAAGGT 7815

Db 7556 CAGCCAGGGCGCGTGCAGACCAACCGTGTCAACATCGCCAAACGAGACTGCATCAAGGT 7615

QY 7816 TGCTGCCATCTTGAACCATGCCCACTACCTAGAGAACCTGCACITTCACCATTTGATGGGT 7875

Db 7616 GCGGCGCGTGTCAACAACCGCTTCTACCTGGAGAACCTGCACITTCACCATCGAGGGCAA 7675

QY 7876 GGATACCCCATTTACTTTGTGAAACCCAGGACCTTCAGAAAGGTGACCTGGCCATCTTGGCCT 7935

Db 7676 GGACACGCACCTACTTCATCAAGACCCACCGCCCGAGAGCGACCTGGGCACGCTGCGGTT 7735

QY 7936 CAGTGGGGGGCGGCAACCCCTGGAGAAATGGGGTCAACGTCACTGTGTCCCAGATCAACAC 7995

Db 7736 GACCAGCGGCCGCAAGGCGCTGGAGAACCGGCATCAACGTGACGGTGTGCGAGTCCACCAC 7795



Db 1066 GA-----ATGGAAGAAGTGAATTCTGATACCATGCCAACAACAACTGTGTCTATTACCTTC 1119

Qy 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGGAACCAACAGAGGAAAGCC 1296

Db 1120 TGGAGACA-----ATGGAAAATT 1137

Qy 1297 CAGTAGTTCTTTCCAGAGGACAGTTTTCATAGATTTCTGGAGAAAATTGATGTGGGAAGCG 1356

Db 1138 AGGTGGATTACGCAAGAAAATAACCCATAGATTTCCGAGAACTTGATATTGGCCGAAG 1197

Qy 1357 AGCCTCCAGAAAGATTCCCTCCGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416

Db 1198 AGCAATTCAAGAGATTCCCTCCGGGATCTTCTGGAGATCACAGCTCTTTCATHGATCAGCC 1257

Qy 1417 TGTGCATCTGAAATTCAATGTGTCTCTGGGAAAGGCAGCCCTGGTTGGCATTATGGCAG 1476

Db 1258 ACAGTTTCTTAAATTCAATATCTCTCTTCAGAAAGGATGCATTTGATGGAGTATATGGCCG 1317

Qy 1477 AAAAGGCTCCCTCCCTTCACATACACAGTTTGACTTTTGTGGAGCTGCTGGATGGCAGGAG 1536

Db 1318 GAAAGGCTTACCGCTTCCCATACTCAGTATGACTTCGTGGAGCTCCTGGATGGCAGCAG 1377

Qy 1537 GCTCCTAACCCAGGAGGCGGAGCCTAGAGGGGACCCCGGCCAGTCTCGGGGAACTGT 1596

Db 1378 GCTGATTGCCAGAGAGCAGCGGAGCCTGCTTGAGACGGAGAGCCGGCGGCAGCGCAG 1437

Qy 1597 GCCCCCTCCAGCCATGAGACAGGTTTCATCCAGTATTTGGATTGAGGAATCTGGCACTT 1656

Db 1438 ATCCGTGAGCCTTCATGAGCCGGCTTTATCCAGTACTTGGATTCTGGAATCTGGCATCT 1497

Qy 1657 GGCTTTTACAAATGACGGAAGGAGTCAAGAGTGGTTTCCCTTCTCACCACCTGCCATTGA 1716

Db 1498 GGCTTTTATAATGATGGGAAAATGACAGCAGGTTGCTCTTTAATACCATTTGTTATAGA 1557

Qy 1717 GTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGACTGCATCTCTGGGACCTG 1776

Db 1558 GTCTGTGGTGAATGTCCCCGAAATTGCCATGGAATGGAGAAATGCGTTTCTGGAACTTG 1617

Qy 1777 CCACTGCTTCCGTGGTTCCTGGGCCCGACTGTGGCAGAGGCTCCTGCCCGCTGCTCTG 1836

Db 1618 CCATTGTTTTCCAGGATTTCTGGGTCCGATTGTTCAAGAGCGCCCTGTCCAGTGTATG 1677

Qy 1837 TAGCGGAAATGGCCAAATACATGAAGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGC 1896

Db 1678 TAGTGGCAACGGGCAGTACTCCAAGGCCGCTGCCCTGTGTTTCAGCGGCTGGAAGGCGAC 1737

Qy 1897 TGAGTGCATGTGCCCAACCAACCAAGTGTATCGATGTGGCCTGACGAAACCATGGCACTG 1956

Db 1738 CGAGTGTGATGTGCCGACTACCCAGTGTATGACCCACAGTGTGGGGTCTGTGGGATTTG 1797

Qy 1957 CATCACGGGCACTGCATCTGCAACCTTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGA 2016

Db 1798 TATCATGGGCTCCTGTGCTTGCAACTCAGGATACAAAGGAGAAAGTTGTGAAGAAGCTGA 1857

Qy 2017 CTGCATGGACCCACATGTTTCAGGCCGGGGTGTCTGCGTGAGAGGCGGAATGCCATTGCTT 2076

Db 1858 CTGTATAGACCCCTGGGTGTTCTAATCATGTTGTGTATCCACGGGGAATGTCACTGCAG 1917

Qy 2077 TGTGGGATGGGAGGCACCAACTGCGAGACCCCGGACATGCTTAGACCAAGTGTTC 2136

Db 1918 TCCAGGATGGGAGGTAGCAATTGTGAATACTGAAGACCATGTGTCCAGACCAGTGTCTC 1977

Qy 2137 AGGCCACGGAACCTTCCCTCCGGACACCCGGCTTTTCAGCTGTGACCCCAAGCTGGACTGG 2196

Db 1978 CGGCCACGGAACGTATCTTCAAGAAAGTGGCTCCTGCACGTGTGACCCCTAAGTGGACTGG 2037

Qy 2197 ACAGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTCCGTAGG 2256

Db 2038 CCCAGACTGCTCAACGAAATATGTTCTGTGGACTGTGGCTCACACGGCGTTTGCATGGG 2097

Qy 2257 GGGCACCTGCCGTGCGAGGATGGCTGGATGGGGCAGCCTGCGACCAAGCGGGCCTGCCA 2316

Db 2098 GGGGACGTGTCGCTGTGAAGAAGGCTGGACGGGCCCAACCTGTAATCAGAGAGCCTGCCA 2157

Qy 2317 CCCGCGCTGTGCCGAGCATGGGACCTGCCCGCAGCGGCAAGTGCAGTGCAGCCCTGGCTG 2376

Db 2158 CCCCGCTGTGCCGAGCACGGGACCTGCAAGGATGGCAAGTGTGAATGCAGCCATGGCTG 2217

Qy 2377 GAATGGGAAACACTGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCC 2436

Db 2218 GAATGGAGAGCACTGCACATC-----GAGGGTTGTCC 2250

Qy 2437 TGGGTTGTCAATGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTG 2496

Db 2251 TGGTCTGTGCAACAGCAATGGAAGATGTACCTTGGACCAAAATGGCTGGCATTTGTGTGTG 2310

Qy 2497 CCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCTGCGGTGACAG 2556

Db 2311 CCAGCCTGGATGGAGAGGAGCAGGCTGTGACGTAGCCATGGAGACTCTTTGCACAGATAG 2370

Qy 2557 CAAAGACAATGATGGAGATGGCCTGGTGGACTGCATGGACCTCTGGACATCATCCAGGA 2616

Db 2371 CAAGGACAATGAAGGAGATGGACTCATTTGACTGATGATGCCATGCCATGGAGACTCTTTGCACAGATAG 2430

Qy 2617 CCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCTTAACCTCTGGACATCATCCAGGA 2676

Db 2431 TTCCTGCCAGAATCAGCCCTATTGTGCGGGACTGCCGGATCCTCAGGACATCATTAGCCA 2490

Qy 2677 GACACAGGTCCTGTGTACAGCAGAACCTTACACTCCTTCTATGACCGCATCAAGTTTCCT 2736

Db 2491 AAGCCTTCAATCGCCTTCTCAGCAAGCTGCCAAATCCTTTTATGATCGAATCAGTTTCCT 2550

Qy 2737 CGTGGGCAGGACAGCACCGCATATATCCCCGGGGAGAACCCCTTTGATGGAGGCAATGC 2796

Db 2551 TATAGGATCTGATAGCACCCATGTTTATACCTGGAGAAAGTCTTCAATAAGAGCCTTGC 2610

Qy 2797 TTGTGTTATTCTGTGGCCAAAGTGTGACATCAGATGGAACCCCCCTTGGTTGGTGTGAACAT 2856

Db 2611 ATCTGTCAATCAGAGGCCAAGTACTGACTGCTGATGGAACCTCCACTTATTTGGAGTAAATGT 2670

Qy 2857 CAGTTTGTCAATAAACCTCTCTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTGA 2916

Db 2671 CTCGTTTTTCCATTACCCAGAAATATGGATATACTATTACCCGCCAGGACGGAATGTTTGA 2730

Qy 2917 CTTGGTGACAAATGGCGGCATCTCCATCATCTCGGTTTCGAGCGGGCACCTTTTCATCAC 2976

Db 2731 CTTGGTGGCAATGGTGGGGCCTCTCTAACTTTGGTATTGTAACGATCCCCATTTCTCTCAC 2790

Qy 2977 ACAGGAGCACACCTGTGGCTGCCATGGGATCGCTTCTTGTGATGGAACCATCATCAT 3036

Db 2791 TCAGTATCATACTGTGTGGATTCCATGGAATGTCTTTTATGTGATGGAATACCTTAGTCTAT 2850

Qy 3037 GAGACATGAGGAGAATGAGATTTCCAGCTGTGACCTGACCTGAGCAATTTTGCCCGCCCAACCC 3096

Db 2851 GGAGAAAGAGAGAATGACATTTCCAGCTGTGATCTGAGTGGATTTCGTGAGGCCAAATCC 2910

Qy 3097 AGTCGTCTCTCCATCCCACCTGACGTCTCTCGCCAGCTCCTGTGCGAGAGAAAGGCCCAT 3156

Db 2911 CATCATTTGTGTCATCACCTTTATCCACCTTTTTCAGATCTTCTCCTGAAGACAGTCCCAT 2970

Qy 3157 TGTCCCGGAAATTTCAGGCTTTGACGAGGAGGAATCTCTATCTCTGCTGCAAGATGAGGCT 3216

Db 2971 CATTTCCGAAACACAGGTACTCCACGAGGAAACTACAATTTCCAGGAACAGATTTGAAACT 3030

Qy 3217 GAGCTACCTGAGCAGCCGGACCCCTGGCTACAAATCTGTCTGAGGATCAGCCTCACCCA 3276

Db 3031 CTCCTACTTGAGITTCAGAGCTGCAGGGTATAAGTCAGTTCTCAAGATCACCATGACCCA 3090

Qy 3277 CCCGACCATCCCTTCAACCTCATGAAGGTGACCTCATGGTAGCGGTGAGGGCCGCT 3336

Db 3091 GTCTATTATTCCATTAAATTTAATGAAGGTTTCATTTATGGTAGCTGTAGTAGGAAGACT 3150

Qy 3337 CTTCAGGAAGTGGTTTCGCTGCAGCCCCCAGACTGTCTCTATTATTTTCAATTTGGGACAAGAC 3396

Db 3151 CTTCCAAAAGTGGTTTCTCGCTTCCCTCACCAAACTTGGCCTATACCTTTTCATATGGGATAAAAC 3210

QY 3397 AGACGTCTACAACCAAGAGGTGTTGGGCTTTTCAGAAAGCCTTTGTTTCCGTGGGTATGA 3456  
Db 3211 AGATGCATATAATCAGAAAGTCTATGTTCTATCTGAAGTGTGTCAGTTGGATATGA 3270  
QY 3457 ATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAAACAACAGTGTCTGAGGGCTA 3516  
Db 3271 GTATGAGTCGTGTTGGACCTGACTCTGTGGGAAAAAGAGGACTGCCATTCTGAGGGCTA 3330  
QY 3517 TGAATTTGACCGCTCCAAGCTTTGGAGGATGGAGCCCTAGACAAACATCATGCCCTCAACAT 3576  
Db 3331 TGAATTTGATCGCTCCAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTGGATGT 3390  
QY 3577 TCAAGTGGTATCCTGCAACAAAGGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCC 3636  
Db 3391 ACAGAACGGTATCTGTACAAAGGAAACGGGGAAACCAAGTTTCATCTCCACAGCCTCC 3450  
QY 3637 TGTATTTGGAGCATCATGGGCAATGGGCGCCCGGAGAACATCTCCTGCCCCAGCTGCAA 3696  
Db 3451 AGTCGTAGTAGCATCATGGGCAATGGGCGAAAGCGCAGCATTTCTGCCCCAGTTGCAA 3510  
QY 3697 CGGCTTGTGACCGGCAACAGCTCCTGGCCCCAGTGGCCCCCTCACCTGTGGCTCTGACGG 3756  
Db 3511 TGGTCAAGCTGATGGTAACAAGTTACTTGGCCCCAGTGGCGTAGCTTGTGGGATCGATGG 3570  
QY 3757 GAGCCTCTATGTGGGTGATTTCAACTACATTTAGAAAGATCTTCCCTCTGGAAATGTCA 3816  
Db 3571 CAGTCTGTACGTAGCGGATTTCAACTACGTGCGGCGGATATTCCTTCTGGAAATGTAAC 3630  
QY 3817 CAACATCTTAGAGCTGAGGAAATAAGATTTTCAGACATAGTCACAGTCCAGCACACAATA 3876  
Db 3631 AAGTGTCTTAGAACTAAGAAATAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATA 3690  
QY 3877 CTACCTGGCCACAGACCCCATGAGTGGGCGGCTCTTCTTGTACAGCAACAGCCGGCG 3936  
Db 3691 CTACCTTGCAACGGATCCAGTCACGGGAGATCTGTACGTTTCTGACACAAACACCCGCG 3750  
QY 3937 GGTCTTTAAATCAAGTCCACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGT 3996  
Db 3751 AATTATCGCCCAAGTCACTTACGGGGGCAAAAGACTTGACTAAATAATGCAGAATCGT 3810  
QY 3997 TGCGGGGACAGGTGACCAAGTCCCTCCCTTTGATGACACTCGCTGCGGGGATGTTGGAA 4056  
Db 3811 CGCAGGACAGGGAGCAATGCCTTCCGTTTTCAGAGGGCAGATGTGGGATGGAGGAA 3870  
QY 4057 GGCCACAGAAGCCACACTCACCAATCCCAAGGGGTATTACAGTGGACAAAGTTTGGCTGAT 4116  
Db 3871 GGCCGTGGAAGCCACACTCATGATGCCAAAGGAATGGCAGTTGTATAAGAAATGATTAAT 3930  
QY 4117 CTACTTCGTGGATGGCAACCATGATCAGACGCGATCGATCAGAATGGGATCATCTCCACCT 4176  
Db 3931 CTACTTTGTTGATGGAACCATGATTAGGAAAGTTGACCAAAATGGAATCATATCAACTCT 3990  
QY 4177 GCTCGGCTCTAATGATCTACATCAGCCCGGCCACTCAGCTGTGATTCTGTCTATGGATAT 4236  
Db 3991 TCTGGGCTCTAACGATTGACTTCAGCCAGACCTTTAACTTTGTGACACCAGCATGCACAT 4050  
QY 4237 TTCCAGGTAAGACTGGAGTGGCCCAACAGACTTAGCCATCAACCCAAATGGACAACACTCACT 4296  
Db 4051 CAGCCAGGTAAGTCTGGAATGGCCCACTGACCTAGCCATTAAACCCCTATGGATAACTCCAT 4110  
QY 4297 TTATGTCTCGACAACAATGTGGTCTGTGAAATCTCTGAAATCTTGAAACCCAGGTGCGCATTTGT 4356  
Db 4111 TTATGTCTCGGATAAATGATGTTTACAGATCACTGAAATCTGAAATCTGCAAGTTTCGCAATTGC 4170  
QY 4357 CGCCGGAGGCCCATGCACTGCCAGGTCCCTGGCATTTGACCACTTCTGCTAAGCAAGGT 4416  
Db 4171 TGCTGGACGGCCCATGCACTGTTCAGGTTCCCGGAGTGG---AATATCCTGTGGGAAGCA 4227  
QY 4417 GGCCATCCACGCAACCTGGAGTCAGCCACCGCTTTGGCTGTTTCACACAATGGGGTCCCT 4476  
Db 4228 CGCGGTGCAGACAACACTGGAATCAGCCACTGCCATTGTGTGTCTACAGTGGGGTCCCT 4287  
QY 4477 GTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCGAGGTCAACCACTAGTGG 4536

Db 4288 GTACATTACTGAAACTGATGAGAAAGAAAATTAACCGGATAAGGCGAGGTCAACAACAGATGG 4347  
QY 4537 AGAGATCTCACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTG 4596  
Db 4348 AGAAATCTCTTAGTGGCCGGAATACCTTCAGAGTGTGACTGCAAAAATGATGCCAACTG 4407  
QY 4597 TGATTTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTC 4656  
Db 4408 TGACTGTTACCAGAGTGGAGATGGCTACGCCAAGGATGCCAAACTCAGTGGCCCATCCTC 4467  
QY 4657 CTTGGCTGTGTGTGCTGATGGGAGCTCTACGTGCGGACCTTGGGAACATCCGAATTCG 4716  
Db 4468 CCTGGCTGCTTCTCCAGATGGTACACTGTATATTGAGATCTAGGAAATATCCGGATCCG 4527  
QY 4717 GTTTATCCGGAAGAAACAAGCCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTCACC 4776  
Db 4528 GGCTGTGTCAAAGAAATAAGCCTTTACTTTAACTCTATGAACCTTCTATGAAGTTGCGTCTCC 4587  
QY 4777 AATTGACGAGGACTCTATCTGTGTTGATACCAACCCGCAAGCACCTGTACACCCAAAAGCCT 4836  
Db 4588 AACTGATCAAGAACTCTACATCTTTGACATCAATGGTACTCACCAATATATACTGTAAAGTTT 4647  
QY 4837 GCCACAGGAGACTACCTGTACAACTTCACTACACTGGGACGGCGACATCACACTCAT 4896  
Db 4648 AGTCACTGGTGATTACCTTTTACAATTTTAGCTACAGCAATGACAAATGATATACTGCTGT 4707  
QY 4897 CACAGACAACAATGGCAACATGGTAAATGTCCGCGGAGACTCTACTGGGATGCCCTCTG 4956  
Db 4708 GACAGACAGCAATGGCAACACCTTTAGAATTAAGCGGACCCAAATCGCATGCCAGTTCCG 4767  
QY 4957 GCTGGTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCACCTCAA 5016  
Db 4768 AGTGGTGTCTCCTGATAACCAAGTGATATGTTTGACAAATAGGAACAAATGGATGTTTGA 4827  
QY 5017 GAGTGTGACCACACAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCT 5076  
Db 4828 AGGCATGACTGTCAAGGACTGGAATTAGTTTGTGTTTACTTACCAATGGCAATAGTGGCCT 4887  
QY 5077 TCTGGCAACCAAAAGCAATGAAAACGGATGGACAAACATTTTATGATAGCACAGCTTTGG 5136  
Db 4888 TTTAGCCACTAAAAGTGATGAAACTGGATGGACAAACGTTTTTTGACTATGACAGTGAAGG 4947  
QY 5137 CCGCTTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGA 5196  
Db 4948 TCGTCTGACAAATGTACGTTTCCAACTGGAGTGGTCAAAACCTGCATGGGACATGGA 5007  
QY 5197 CAGTTTCACTGATGTCCAGGTAGAGACCTCCAGC--AAGGATGATGTCAACCATAACCC 5253  
Db 5008 CAAGGCTATCAGAGTGGACATTTGAGTCTATAGCCGAGAGAAGATGTGACAGTCACTTC 5067  
QY 5254 CAACCTGTCTGCCTCAGGGCCTTCTACACACTGCTGCAAGACCAAGTCCCGAACAGCTA 5313  
Db 5068 AAATCTGTCTCGATCGATTCTTTCTACACCATGTTCAAGATCAGTTAAGAAACAGCTA 5127  
QY 5314 CTACATCGGGGCCGATGGCTCCTTGGCGCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCT 5373  
Db 5128 CCAGATTGTTTACGGCTCCTCTCAGAAATTAATAACCGCAGTGGCTGGACTCACACTA 5187  
QY 5374 GCAGACTGAGCCCACTTGTGGCTGGACCGTCAACCCCAACCGTGGCAAGAGGAATGT 5433  
Db 5188 CCAACAGAGCCGACGTTCTGGCTGGACCGCTAATCCGACGGTTGCCAAAAGAAACAT 5247  
QY 5434 CACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGC 5493  
Db 5248 GACTTTGCTGTGGAGAACCGTCAAAACTTGGTGAATGGAGATTCCGAAAAGAGCAAGC 5307  
QY 5494 TCGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGGTGCACAACCGAAATCTCCTATC 5553  
Db 5308 CCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCTTTC 5367  
QY 5554 TCTGGACTTTTGTATCGCGTAAACACGACAGAGAAAGATCTATGATGACCAACCGCAAGTTTCAC 5613

Db 5368 AGTTGACTTTGATCGAACAAACAGACAGAAAAGATCTATGACGACCACCGTAAATTTCT 5427

QY 5614 CCTTCGGATTCTGTACGACACGCGGGGGCCAGCCTCTGGTCACCCCAGCAGGCT 5673

Db 5428 ACTGAGGATCGCCTACGACACGCTGTGGCACCCGACTCTCTGGTGCCAAAGCAGCAAGCT 5487

QY 5674 GAATGGTGTCAACGTGACATACTCCCTGGGGTTTACATGTGCTGGCATCCAGAGGGGCAT 5733

Db 5488 GATGGCCGTCAATGTACCTATTATCATCCACAGGTCAAATGCCAGCATCCAGCGAGGCAC 5547

QY 5734 CATGTCTGAAAGAAATGGAATACGACCAGGCGGGCCGCATCACATCCAGGATCTTCGCTGA 5793

Db 5548 NACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTTGCTGA 5607

QY 5794 TGGGAGACATGGAGCTACACATACTTAGAGAAGTCCATGGTCTGCTACTACACAGCCA 5853

Db 5608 TGGTAAACATGGAGTTACACATATTAGAAAAGTCCATGGTTCTTCTGCTTCATAGCCA 5667

QY 5854 GAGGCAGTATATCTTTGAGTTCGACAAGAAATGACCGCCCTCTCTTCTGTGACGATGCCCAA 5913

Db 5668 GCGGCAGTACATCTTCGAATACGATATGTGGGACCGCCTGTCTGCCATCACCATGCCCAG 5727

QY 5914 CGTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCA 5973

Db 5728 TGTGGCTCGCCACACCATGCAGACCATCCGATCCATTGGGTACTACCGCAACATATACAA 5787

QY 5974 GCCCCCTGAGGGCAATGCCTCAGTCAATACAGGACTTCACTGAGGATGGGCACCTCTCTCA 6033

Db 5788 CCCCCCGGAAAGCAACGCCTCCATCATCACGGACTACAAAGAGGAGGGCTGCTTCTTACA 5847

QY 6034 CACCTTCTACCTGGGCACCTGGCCGAGGGTGATATACAAGTATGGCAAACTGTCAAAGCT 6093

Db 5848 AACAGCTTCTTGGGTACAAAGTCGAGGGTCTTATTCAAATACAGAAGGCAGACTAGGCT 5907

QY 6094 GGCAGAGACGCTCTATGACACCAACCAAGGTCAAGTTTCACTATGACGAGACGGCAGGCAT 6153

Db 5908 CTCAGAAATTTATATGATAGCAACAGAGTCAGTTTTTACCTATGATGAAACAGCAGAGT 5967

QY 6154 GCTGAGACCATCAACCTACAGAAATGAGGGCTTCACTGACCATCCGCTACCGTCAGAT 6213

Db 5968 CCTAAAGACAGTAAACCTCCAGAGTGATGGTTTTATTGCAACCATTAGATACAGGCAAAAT 6027

QY 6214 TGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCCCG 6273

Db 6028 TGGTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGGTAAATGCAAG 6087

QY 6274 TTTTGACTACAACTATGACAAACAGCTTCCGGGTGACCAAGCATGCAGGCTGTGATCAACGA 6333

Db 6088 ATTTGACTATAGCTATGACAAACAGCTTTCGAGTGACCAAGCATGCAGGGTGTGATCAATGA 6147

QY 6334 GACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTT 6393

Db 6148 AACGCCACTGCCATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAGTT 6207

QY 6394 TGGGAAGTTTGGTGTCAATTTACTATGACATTAACCAAGATCATCACCAAGCTGTCTGAC 6453

Db 6208 TGGAAAGTTTGGAGTTATATATTTATGATATTAACCAAGATCAATTTCTACAGCTGTAATGAC 6267

QY 6454 CCACACCAAGCATTTTGTATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCG 6513

Db 6268 CTATACGAAGCACTTTGTATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTCAG 6327

QY 6514 CTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGCGGAGTAGTGAAGAGGA 6573

Db 6328 GTCGCTCATGTACTGGATTACAATTCAGTATGATAACATGGGTCCGGTAAACCAAGAGAGA 6387

QY 6574 GCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGATGCTGACGG 6633

Db 6388 GATTAAATAGGGCCCTTTGCCAACACCAACCAATATGCTTATGAATATGATGTTGATGG 6447

QY 6634 CCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGGTACAGCTACGACCTCAA 6693

Db 6448 ACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTGAA 6507

QY 6694 TGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACCTACGGTATGA 6753

Db 6508 TGGAAACCTCCATTTACTGAACCCCAAGTAACAGTGCAGTCTGACACCCCTTCGCTATGA 6567

QY 6754 CATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGCTT 6813

Db 6568 CCTGCGAGACAGAATCACTCGACTGGTGATGTTCAATATATCGTTGGATGAAGATGGTTT 6627

QY 6814 CCTGAGGACCGGGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCCTGCTCATCAAGGC 6873

Db 6628 CCTACGTCAAAGGGGCACGGAATCTTTGAATATAGCTCCAAGGGCTTCTAACTCGAGT 6687

QY 6874 CTACAAACCGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCTGGGGCGGCGCT 6933

Db 6688 TTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCTGGGAAGGCGTGT 6747

QY 6934 GTCCAGCAAGACAGCCACAGCCACCACTCCAGTCTTCTTATGACAGACCTGACCAACCC 6993

Db 6748 TTCTAGCAAAACCAAGTCTAGGACAGCACCTGCAGTTTTTTTTATGCTGACTTAACCTTATCC 6807

QY 6994 CACCAAGGTCACCCACCTGTATACAACTCCAGCTCTGAGATCACTCCCTCTACTACGA 7053

Db 6808 CACTAGGATTACTCATGTCTACAAACCATTCGAGTTTCAGAAATACCTCCCTGTATTATGA 6867

QY 7054 CTTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGATGAGTTTTACATAGCTTG 7113

Db 6868 TCTCCAAGGACATCTTTTGCCATGGAAATCAGCAGTGGGATGAATTTATATTGCATC 6927

QY 7114 TGACAACATCGGGACCCCTCTTGTCTCTTTAGTGGAAACAGGTTTGATGATCAAGCAAAAT 7173

Db 6928 GGATAACACAGGGACACCACTGGCTGTGTTCACTAGTAGCAATGGGCTTATGCTGAAACAGAT 6987

QY 7174 CCTGTACACAGCCTATGGGGAGATCTACATGGATACCAACCCCACTTTCAGATCATCAT 7233

Db 6988 TCAGTACACTGCATATGGGAAATCTATTTTGACTCTAATATATGACTTCAACTGGTAAT 7047

QY 7234 AGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGCGCGGAGA 7293

Db 7048 TGGATTTCAITGGTGGCCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAAAGAGA 7107

QY 7294 TTATGATGTGCTGGCCGGACGCTGGACTAGCCACGACCCAGCTGTGGAAGCACCTTAG 7353

Db 7108 TTATGACATTTTGGCAGGACGGTGGACAAACACCTGACATAGAAATCTGGA---AAAGAAT 7164

QY 7354 TAGCAGCAACGTTCATGCCTTTTAACTCTATATGTTCAAAAACAACACCCCATCAGCAA 7413

Db 7165 TGGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCTTCAAGCAA 7224

QY 7414 CTCCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGATT 7473

Db 7225 AATCCATGACGTGAAAGATTACATCACAGATGTTAACAGCTGGCTGGTGACATTTGGTTT 7284

QY 7474 CCAGCTACAAACGTGATCCCTGGTTATCCCAAAACAGACATGATGCCATGGAACCCCTC 7533

Db 7285 CCATCTGCACAAATGCTATTCTCTGGATTCCCTGTTCCCAAATTTGATTTAACAGAACCTTC 7344

QY 7534 CTACGAGCTCATCCACACACAGATGAAACGACGAGTGGGACAAACAGCAAGTCTATCCT 7593

Db 7345 TTACGA-----ACTTGTGAAGATCAGCAGTGGGATGATATACCGCCCATCTT 7392

QY 7594 CGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTTGA 7653

Db 7393 CGGAGTCCAGCAGCAAGTGGCGCGCAGGCCAAAGCCCTTCCCTGTCGCTGGGGAAGATGGC 7452

QY 7654 CCAGCTCTATGGCTCCACAATCACCAAGCTGCCAGCAGGCTCCAAAGACCAAGAAAGTTTGC 7713

Db 7453 CGAGGTGCAGGTGAGCCGGCGCCGGCCCGC---GGCGCGCAGTCCCTGGCTGTGGTTTCGC 7509

QY 7714 ATCCAGCGGCTCAGTCTTTTGGCAAGGGGGTCAAGTTTTCCTTTGAAGGATGGCCGAGTGAC 7773

Db 7510 CACGGTCAAGTCCCTGATCGGCAAGGGCGCTCATGCTGGCCGTGAGCCAGGCGCGGTGCA 7569



QY	272	CTGCGGGAGCTGGGGCTGGAAGAAGTAAACGCCCCCTCACGGGACCCTGTACCGACAGAC	331
Db	520	CTTTGCCGAACCTGGGCATCTGTGA--GCCCTCCCACACCGAAGCGGTACTGTCTCGAC	576
QY	332	ATTGGCCCTGCCCAATATGCGGCTACTCCATGGSGGCTGGCTCTGATGCCGACATGGAGGCT	391
Db	577	ATGGGGATCCTTACCAGGGCTACTCCCTTAGACAGGGTCTGACGCCGACTCCGACACC	636
QY	392	GACACGGTGCTGTCCCTTGAGCACCCCGTGCCTGTGTGGGCGCGGAGCACACGGTCAGGG	451
Db	637	GAGGGAGGGATGTCTCCAGAACACGCCATCAGACTGTGGGCAGAGGGATAAAATCCAGG	696
QY	452	CGCAGCTCCTGCCTGTCCAGCCGGGCCAAATTCCAATCTCACACTCACCCGACACCGAGCAT	511
Db	697	CGCAGTTCCGGCCTGTCCAGTCTGTAATACTCGGCCCTTACCTGACTGACTCTGACAAC	756
QY	512	GAACAACACTGAGACTGATCATCCGGGCGGCC----TGCAGAACCAACGCGCGGCTCCGGAC	567
Db	757	GAACAACAATCAGATGATGAGAACGGTCTGCTCCATTCACCTACATCCTCGCCTAGTCTC	816
QY	568	GCCGCCGCCGCGCTCTCGCACGCCCAACACCCCAACCCAGCACCAACGCGGCTCCATTAA	627
Db	817	CTCCCATCTGCTCAGCTGCCTAGCTCCCATTAATCTCCACCAGTTAGTGGC-AGATGCC	875
QY	628	CTCCCTGAACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCACCGACCA	687
Db	876	ATTGCTAGACAGCAACACCTCCCATCAAATCATGACACCAACCCCTGATGAGGAATCTC	935
QY	688	CTCGCTCTCCGAGAGCCCCCTTCCGGCGGCGGCCAGGAGCCTGCCCCACGCCCAGGAGAA	747
Db	936	CCCCAATTATACCTGCTCAGAGCATGCTCAGGCCCCAGCAAGCCTCCAGCAGTGGCCC	995
QY	748	CTGGCTGCTCAACAGACAACATCCCCCTGGAGACAGGAACCTAGCAAGCAGCATTCTCT	807
Db	996	TCCGAACCAACACAGCCAGTCGACTCTGAGGCC-----CCTCTCCC	1037
QY	808	AGGACATTGCAGGACAACTCATTTAGATGGACATTTCTGGCGCCTCCCGGCATGATGG	867
Db	1038	ACCCCTCACAACACACAGCTGTCCCATCACCTCGTCCGCCAACTCCCTCAACAGGAA	1097
QY	868	GGCTTACAGTGACGGGCACTTCTCTTCAAGCTGGAGGCACCTCCCGCTCTTCTGCAC	927
Db	1098	CTCACTGACCAATCGGCGGAGTCAGATCCACGCCCGGCCCGCCCAATGACCTGGC	1157
QY	928	CACATCACAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCT	987
Db	1158	CACCACACCAGAGTCCGTTCACTTCAGGACAGCTGGGTGTAAACAGCAACGTGCCACT	1217
QY	988	GCCCCGCAGCACCTTCCGCCCGCGGCCCTTTAACCTCAAGAAGCCCTCCAGTACTGTAA	1047
Db	1218	GGAGACCCGGCATTCTCTTCAAGACCTCTCTCGGGAGCACACCT-----T	1265
QY	1048	CTGGAAGTCCGCAGCCCTGAGCGCCATCGTCTATCTCAGCCACTCTGGTCTATCTGCTGGC	1107
Db	1266	GTTTCAGCAGCTCTTCCCGGGATACCTTTGACCTCAGGAACGGTTTACACGCCCGCC	1325
QY	1108	ATACTTTGTGGCCATGCACCTGTTTGGCCATAACTGGCACCTGCAGCCGATGGAGGGCA	1167
Db	1326	CGCCTGCTGCCAGGAATACTTTCTCCA-----GGAGGCTTTCAA	1367
QY	1168	GATGTATGAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCC-C	1226
Db	1368	GCTGAAGAAGCCCTCCAAATACTGCAGCTGGAATGTGCTGCCCTCTCCGCCCATGCGCG	1427
QY	1227	TATACCCCTCAGGGGCACCTGGCTTAGAGACCCCTGACAGGAAGGCAAGGAACACAG	1286
Db	1428	GGCCCTCTCTTGGCTATTTTGTGGCGTATTTTCATAGTGGCCTGGTCGTTGAAAAACAG	1487
QY	1287	AAGGAAAGCCAGTAGTTTCTTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATG	1346
Db	1488	CAG-----CATAGACAGTGGTGAAGCAGAAG	1513

QY	1347	TGGAAGGCGAGACCTCCAGAAAGATTCTCTCTGGCACATTTCTGGAGATCTCAAGTGTTC	1406
Db	1514	TTGGTCGGCGGTAAACACAAGAAGTCCACACAGGGGTGTTTGGAGGTACAAAATTCACA	1573
QY	1407	TAGACCATCTGTGTCATCTGAAATTCATATGTGTCTCTGGGAAAGGACGCCCTGGTTGGCA	1466
Db	1574	TCAGTCAGCCCGAGTCTTAAAGTTCAACATCTCCCTCGGAAGGACGCTCTCTTTGGTG	1633
QY	1467	TTTATGGCAGAAAAGCCCTCCCTCCCTTACATACACAGTTTGACTTTGTGGAGCTGCTGG	1526
Db	1634	TTTACATAAGAAGAGGACTTCCACCATCTCATGCCAGTATGACTTTCATGGAACGTCTGG	1693
QY	1527	ATGGCAGGAGGTCTTAACCCAGGAGGCGGAGCCCTAGAGGGGACCCCGGCCAGTCTC	1586
Db	1694	ACGGGA-----AGGAGAAAGTGGAGTGTGGTTGAGTCTCCAGGGAACGCC	1738
QY	1587	GGGGAACCTGTGCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTCAGGAA	1646
Db	1739	GGAGCATACAGACCTTGGTTTCAGAAATGAAGCCGTGTTGTGCAGTACCTGGATGTGGCC	1798
QY	1647	TCTGGCACTTGCTTTTACAATGACGGAAGGAGTCAGAAAGTGGTTTCTCTCTCACCA	1706
Db	1799	TGTGGCATCTGCCCTTCTACAATGATGGAAGGACAAAGAGATGGTTTCTTCAATACTG	1858
QY	1707	CTGCCATTGAGTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGCTGACTGCATCT	1766
Db	1859	TTGTCTTAGATTCACTGTCAGGACTGTCCACGTAACTGCCATGGGAATGCTGAATGTGT	1918
QY	1767	CTGGGACCTGCGCACTGCTTCTCTGGGTTTCTTGGCCCCGACTGTGGCAGAGCCTCTGCC	1826
Db	1919	CCGGGTGTGTCACTGTTTCCACGGATTCTTAGGAGCAGACTGTGCTAAAGTGCCTGCC	1978
QY	1827	CCGTGCTCTGTAGCGGAAATGGCCCAATACATGAAAGGCAGATGCTTGTGCCACAGTGGCT	1886
Db	1979	CTGTCTGTGTCAGTGGGAATGGACAATATTTCTAAAGGACGTCGCCAGTGTACAGCGCT	2038
QY	1887	GGAAAGGCGCTGAGTCGGATGTGCCCAACCAACCAAGTGTATCGATGTGGCTGCAGCAACC	1946
Db	2039	GGAAAGGTGCAGATGCGACGTGCCCATGAATCAGTGCAATCGATCCCTTCGCGGGGCC	2098
QY	1947	ATGGCACCTGCATCACGGGCACCTGCATCTGCAACCCCTGGCTACAAGGGCAGAGCTGTG	2006
Db	2099	ACGGCTCCTGCATTGATGGGAATGTGTCTGCTCTGTGGCTACAAAGGCGAGCACTGTG	2158
QY	2007	AGGAAGTGGACTGCATGGACCCCAACATGTTTCAGGCCGGGGTGTCTGCGTAGAGGCGAAT	2066
Db	2159	AGGAAGTTGATTGCTTGGATCCCACTGCTCCAGCCACGGAGTCTGTGTGAATGGAGAAT	2218
QY	2067	GCCATTGCTTTGTGGGATGGGAGGCACCAACTGCGAGACCCCGAGGCTTGCATGCTTAG	2126
Db	2219	GCCTGTGCAGCCCTGGCTGGGTGGTCTGAACTGTGAGCTGGCGAGGTCCAGTGCACG	2278
QY	2127	ACCAGTGTTCAGGSCACGGAACTTCTCTCCCGACACCGGGCTTTGCAGCTGTGACCCAA	2186
Db	2279	ACCAGTGCAGTGGGCATGGCAGTACCTGCTGCACACGGGCCCTCTGCAGCTGCGATCCCA	2338
QY	2187	GCTGGACTGGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGSCATGGCG	2246
Db	2339	ACTGGATGGGTCCCGACTGCTCTGTGTGAAGTGTGCTCAGTAGACTGTGGCACTCAGCGG	2398
QY	2247	TGTGCGTAGGGGSCACCTGCCGCTGCGAGGATGGCTGGATGGGGGCAGCCTCGACCCAGC	2306
Db	2399	TCTGCATCGGGGAGCCTGCCGCTGTGAAGAGGGCTGGACAGGCGCAGCGTGTGACCAGC	2458
QY	2307	GGGCCTGCCACCCCGCTGTGCCGAGCATGGGACCTGCCGCGACGGCAAGTGCAGTGCA	2366
Db	2459	GCGTGTGCCACCCCGCTGCATTGAGCACGGGACCTGTAAAGATGGCAATGTGAATGCC	2518
QY	2367	GCCCTGGCTGGAATGGCGAACAACCTGCACCACTGCTCACTATCTGGATAGGGTAGTTAAAG	2426
Db	2519	GAGAGGGCTGGAAATGGTGAAACACTGCACCAATGGTAGGCCAAACCGCAGGCCACCGAAACAG	2578
QY	2427	AGGGTTGCCCTGGGTGTGCAATGGCAACGGCAGATGTACTTTAGACCTGAAATGGTTGGC	2486

Db 2579 ATGGTGCCTCGACTTGTGCAACGGTAACGGGAGATGCACACTGGGTGACAAACAGCTGGC 2638  
Qy 2487 ACTGCGTCTGCCAGCTGGGTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCCT 2546  
Db 2639 AGTGTGTCTGCCAGACCGGCTGGAGAGGGCCCGGATGCAACGTTGCCATGGAAACTTCCT 2698  
Qy 2547 GCGGTGACAGCAAAGACAAATGATGGAGATGGCCCTGGTGGACTGCATGGACCTGACTGCT 2606  
Db 2699 GTGCTGATAACAAGGATAATGAGGGAGATGGCCCTGGTGGATTGTTGGACCTGACTGCT 2758  
Qy 2607 GCCTCCAGCCCTGTGCCCATATCAACCCGCTGTGCTTGGCTCCCTTAACCCCTCTGGACA 2666  
Db 2759 GCCTGAGTCAGCCTGTGCAAAACAGCCTGTCTGCGCGGGGTCCCGGACCCACTGGACA 2818  
Qy 2667 TCATCCAGGAGACACAGGTCCCTGTGTACAGCAAAACCTACACTCCTTCTATGACCGCA 2726  
Db 2819 TCATTGAG-----CAGGGCCAGACGGATTGGCCCGCAGTGAAGTCCCTTCTATGACCGTA 2872  
Qy 2727 TCAAGTTCCCTCGTGGGCAGGGACAGCACGCACATAAATCCCGGGAGAAACCCCTTTGATG 2786  
Db 2873 TCAAGCTCTTGGCAGCAAGGATAGCACCCACATCATCTCTGGAGAGAACCTTTCAACA 2932  
Qy 2787 GAGGGCATGCTTGTGTTATTCGTGGCCAAAGTATGACATCAGATGGAACCCCTTGGTTG 2846  
Db 2933 GCAGCTTGGTTTCTCTCATCCGAGGCCAAGTAGTAACATACAGATGGAATCCCTGGTCG 2992  
Qy 2847 GTGTGAACATCAGTTTTGTCAATAACCCCTCTCTTTGGATATACAAATCAGCAGGCAAGATG 2906  
Db 2993 GTGTGAACGTGTCTTTTGTCAAGTACCCAAATACCGCTACACCATCACCCGCCAGGATG 3052  
Qy 2907 GCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCCTGCGGTTTCGAGCGGCAC 2966  
Db 3053 GCAGCTTCGACCTGATCGAAATGGAGGTGCTTCTTGACTCTACACTTTGAGCGAGCCC 3112  
Qy 2967 CTTTCATCACACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTTCTTGTCTATGAAA 3026  
Db 3113 CTTTCATGAGCCAGGAGCCACTGTGTGGCTGCGTGGAAACAGCTTTTACGCCATGGACA 3172  
Qy 3027 CCATCATCATGAGACATGAGGAGAAATGAGATTCCAGCTGTGACCTGAGCAATTTTGGCC 3086  
Db 3173 CCCTGGTGATGAAGACCGAGGAGAACTCCATCCCCAGCTGTGACCTCAGTGGCTTTGTCC 3232  
Qy 3087 GCCCAAACCCAGTCGTCTCTCCATCCCACTGACGTCCTTCCGCAAGCTCCTGTGCAGAGA 3146  
Db 3233 GGCCTGATCCAATCATCTCTCCCTCCACTGTCCACCTTCTTTAGTGTGCTGCCCTGGGC 3292  
Qy 3147 AAGGCCCCATTGTGCCGAAATTCAGGCTTTGAGGAGGAAATCTCTATCTCTGGCTGCA 3206  
Db 3293 AGAATCCCATCGTGCCTGAGACCCAGGTTCTTCATGAAGAAATCGAGCTCCCTGGTTCCA 3352  
Qy 3207 AGATGAGGCTGAGCTACCTGAGCAGCCGACCCCTGGCTACAAATCTGTCCTGAGGATCA 3266  
Db 3353 ATGTGAACCTTCGCTATCTGAGCTCTAGAACTGACAGGTTACAAGTCACTGCTGAAGATCA 3412  
Qy 3267 GCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGGTAGCGGTGG 3326  
Db 3413 CCATGACCCAGTCCACAGTGCCTGAACTCATTAGGTTTCACTGATGGTGGCTGTGTCG 3472  
Qy 3327 AGGGCCGCTCTTTCAGGAAGTGGTTGCTGACGCCCCAGACCTGTCTCTATTTATTTCAAT 3386  
Db 3473 AGGGGCATCTCTCCAGAAAGTCATTCAGGCTTCTCCAACTGCGCCCTCCACCTTCATCT 3532  
Qy 3387 GGGACAAGACAGACGCTCTACAACCAAGAGGTGTTGGCTTTTCAAGACCTTTGTTTCCG 3446  
Db 3533 GGGACAAGACAGATCGCTATGGCCAAAGGGTGTATGGACTCTCAGATGCTGTGTGTCG 3592  
Qy 3447 TGGGTTATGAATATGAATCTGCCAGATCTAATCTGTGGGAAAAAAGAACACAGTGC 3506  
Db 3593 TCGGGTTGAATATGAGACCTGTCCCAGTCTAATCTCTGGGAGAAAAGACAGCCCTCC 3652  
Qy 3507 TGCAGGGCTATGAAATTGACCGCTCCAAGCTTGGAGGATGGAGCCCTAGACAAACATCATG 3566

Db 3653 TTCAGGGATTGAGCTGGACCCCTCCAACCTCGGTGGTGGTCCCTAGACAAACACCACA 3712  
Qy 3567 CCTCAACATTCAAAGTGGTATCCTGCACAAAAGGGAATGGGAGAACCAAGTTGTGTCTC 3626  
Db 3713 TCCTCAATGTTAAAAGTGAATCCTACACAAAAGGCACTGGGAAAACCAAGTTCTTGACCC 3772  
Qy 3627 AGCAGCCTCCTGTCTATTTGGAGCATCATGGGCAATGGCGCCGAGAGAGCATCTCCTGCC 3686  
Db 3773 AGCAGCCTGCCATCATCACCAGCATCATGGGCAATGGTCCCGCCGAGCATTTCTGTG 3832  
Qy 3687 CCAGCTGCAACCGCCTTGTGACGGCAACAAGTCTTGGCCCGAGTGGCCCTCACCTGTG 3746  
Db 3833 CCAGCTGCAACCGCCTTGTGAGGCAACAAGTCTTGGCCCGAGTGGCTTGGCTGTG 3892  
Qy 3747 GCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTCCCTCTG 3806  
Db 3893 GAATCGATGGGAGCCTCTATGTGGGTGACTTCAATTACATCCGACGATCTTCCCTCTC 3952  
Qy 3807 GAAATGTCAACCAACATCCTAGAGCTGAGGATAAAGATTTTCAGACATAGTCACAGTCCAG 3866  
Db 3953 GAAATGTGACCAAGCATCTTGGAGTTACGAATAAAGATTTTAAACATAGCAACACCCAG 4012  
Qy 3867 CACACAAATACTACTTGGGCCACAGACCCCATGAGTGGGGCGCTCTTCTTTCTGACAGCA 3926  
Db 4013 CACACAAAGTACTACTTGGCAGTGGACCCCGCTGCTCCGGCTCGCTCTACGTGTCGACACCA 4072  
Qy 3927 ACAGCCGCGGGTCTTTAAAATCAAGTCCACTGTGTGGTGAAGGACCTTGTCAAGAACT 3986  
Db 4073 ACAGCAGGAGAACTCTACCGCTCAAGTCTCTGAGTGGAAACCAAGACCTGGCTGGAAAT 4132  
Qy 3987 CTGAGGTGTTGCGGGACAGGTGACCAAGTGCCTCCCTTTGATGACACTCGCTGCGGG 4046  
Db 4133 CGGAAAGTTGTGGCAGGACGGGAGAGCAGTGTACCTTTTATGAAGCCCGCTGCGGG 4192  
Qy 4047 ATGGTGGGAAGGCCACAGAAAGCCACACTCAACAAATCCAGGGGTATTACAGTGGACAA 4106  
Db 4193 ATGGAGGGAAGGCCATAGATGCAACCCCTGATGAGCCCGAGAGGTATTGACAGTAGACAA 4252  
Qy 4107 TTGGCTGATCTACTTCGTGGATGGCACCACTACCAATCCAGGGGTATTACAGTGGACAA 4166  
Db 4253 ATGGGCTCATGTACTTGTGATGCCACCCTATGATCGGAGGTGACCAAGATGGAATCA 4312  
Qy 4167 TCTCCACCCCTGCTCGCTCTAATGATCTCATCATCAGCCCGGCACTCAGCTGTGATTCG 4226  
Db 4313 TCTCCACCCCTGCTGGCTCCAATGACCTCACTGCTCCGCTCCGCGCGCTGAGCTGTGATCCA 4372  
Qy 4227 TCATGGATATTTCCAGGTAAGACTGGAGTGGCCCAACAGACTTAGCCATCAACCCCAATGG 4286  
Db 4373 GCATGGATGTAGCCAGGTTTCGTCTGGAGTGGCCCAACAGACCTTGTGTCAATCCCATGG 4432  
Qy 4287 ACAACTCATTCTTATGCTCGACAAATGTTGTTCTGCAATCTCTGAAAAACCCACAGG 4346  
Db 4433 ATAACCTCTTGTATGTTCTAGAGAACAAATGTCATCTTCGAATCACCGAGAACCCCAAG 4492  
Qy 4347 TGGCATTGTGCGCGGAGGCCCATGCATGCCAGGTCCCTGGCATTTGACCACTTCCCTGC 4406  
Db 4493 TCAGCATCATTTGCGGACGCCCCATGCATGCCAAGTTCTTGGCATTGACTACT---CAC 4549  
Qy 4407 TAAGCAAGGTGGCCATCCACGCAACCTTGAGTCAGCCACCCGCTTTTGGCTGTTTCACACA 4466  
Db 4550 TCAGCAAACTAGCCCATTCACCTCTGCCCTGGAGTCAGCCAGTGCATGCTGCTTCTCACA 4609  
Qy 4467 ATGGGCTCCTGTATATTGCTGAGACTGTAGAAAAAAGATCAACCCCATCAGGCAGGTCA 4526  
Db 4610 CTGGGCTCCTCTACATCACTGAGACAGATGAGAAGAAAGATTAAACCGTCTACGCCAGGTA 4669  
Qy 4527 CCATAGTGGAGAGATCTCACTCGTTTGTGGGGCCCCCAGTGGCTGTGACTGTAAAAATG 4586  
Db 4670 CAACCAACCGGGAGATCTGCCCTTTTAGCTGGGGCAGCCCTCGGACTCGCACTGCAAAAACG 4729  
Qy 4587 ATGCCAACTGTGATGTTTTTCTGGAGACGATGGTTATGCCAAGATGCAAAAGTTAAATA 4646  
Db 4730 ATGTCAATTGCAACTGCTATTTCAGGAGATGATGCCTACGCGACTGATGCCATCTTGAATT 4789

QY	4647	CCCCATCTTCCTTGGCTGTGTCTGTGATGGGGAGCTCTACGTGGCCGACCTTGGGAACA	4706
Db	4790		
QY	4707	TCCGAATTCGGTTTATCCGGAAGAACAAAGCCCTTTCCTCAACACCCAGAACATGTATGAGC	4766
Db	4850		
QY	4767	TGCTTTCACCAANTGACCAGGAGCTCTATCTGTTTGTATACCACCGGCAAGCACCTGTATACA	4826
Db	4910		
QY	4827	CCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTTCACCTACACTGGGGACGGCGACA	4886
Db	4970		
QY	4887	TCACACTCATCACAGACAACAATGGCAACATGGTAAATGTCCGCCGAGACTCTACTGGGA	4946
Db	5030		
QY	4947	TGCCCCCTCTGGTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGACCAACA	5006
Db	5090		
QY	5007	GTGCACTCAAGAGTGTGACCACACAAGGACACGAGTTGGCCATGATGACATACCATGGCA	5066
Db	5150		
QY	5067	ATTCCGGCCTTCTGGCAACCAAAGCAATGAAACGGATGGACAAACATTTTATGAGTACG	5126
Db	5210		
QY	5127	ACAGCTTTGGCCGCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAA	5186
Db	5270		
QY	5187	GTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTCCA--GCAAGGATGATGTCA	5243
Db	5330	GGGAAATGGAGAAATCTATTACCATTTGACATTGAGAACTCCAACCGTGATGATGACGTCA	5389
QY	5244	CCATAACCAACCACTGTCTGCCTCAGGCGCCTTCTACACACTGTCTCAAGACCAAGTCC	5303
Db	5390	CTGTTCATCACCACCTCTCTTCAGTAGAGGCTCCTACACAGTGTGTACAAGATCAAGTTC	5449
QY	5304	GGACAGCTACTACATCGGGGCCGATGGCTCTCTTCGGCTGCTGCTGSCCAACGGCATGG	5363
Db	5450		
QY	5364	AGGTGGCGCTGAGACTGAGCCCCCACTTGTGCTGGTGGCACCGTCAACCCACCGTGGCA	5423
Db	5510		
QY	5424	AGAGGAATGTACGCTGCCCATCGACAAACGGCTCAACCTGGTGGAGTGGCCCGCAGCGCA	5483
Db	5570		
QY	5484	AAGAGCAGGCTCGGGGCGAGGTCACTGTCTTTGGGCGCCGGCTGCGGTGCAACCCGAA	5543
Db	5630	AGGAACAGATTAAAGGCAAAAGTCAACCATCTTTGGCAGGAAGCTCCGGGTCCATGGAAGAA	5689
QY	5544	ATCTCCTATCTCTGACTTTGATCGGTAAACACGCAAGAGAGATCTATGATGACCACC	5603
Db	5690		
QY	5604	GCAAGTTCACCCCTTCGATTCTGTACGACCGCGGGCGGCCAGCCCTCTGTCAACCCA	5663
Db	5750		
QY	5664	GCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGGTTACATTGCTGGCATCC	5723
Db	5810		
		GCAGCGGGCTGGCAGTGTCAACGTGTATCTTCAATGGCGCGCTGGCTGGGCTTC	5869

QY	5724	AGAGGGCATCATGTCTGTAAGAATGGAATACGACAGCGGGCCGCATACATCCAGGA	5783
Db	5870	AGCGTGGGCCATGAGCGAGAGGACAGACATCGACAAAGCAAGCCGCATCGTGTCGGCA	5929
QY	5784	TCTTCGCTGATGGGAAGACATGAGGCTACACATACCTTAGAGAAGTCCATGGTGCTGCTAC	5843
Db	5930	TGTTCCGCTGACGGGAAAGTGTGAGGCTACTCCTACCTTGACAAGTCCAATGGTCTCTCTGC	5989
QY	5844	TACACAGCCAGAGGCAGTATATCTTTGAGTTTCGACAAGAAATGACCGCCTCTCTTCTGTGA	5903
Db	5990	TTACAGAGCCAAAGTCAGTATATATTTGAGTATGACTCCTCTGACCGCCTCCTTGCCGTCA	6049
QY	5904	CGATGCCCAAAGTCGTGGCGGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAA	5963
Db	6050	CCATGCCAGCGTGGCCCGGCACAGCATGTCCACACACACCTCCATCGGCTACATCCGTA	6109
QY	5964	ACATCTATCAGCCCCCTGAGGGCAATGCCCTCAGTCAATACAGGACTTCACTGAGGATGGGC	6023
Db	6110	ATATTACAAACCGCCTGAAAGCAATGCTTCGGTCACTTTGACTACAGTGATGACGGCC	6169
QY	6024	ACCTCCTTACACACCTTCTACCTGGGCACCTGGCCGACGGTGATATACAAGTATGGCAAAAC	6083
Db	6170	GCATCCTGAAGACCTCCTTTTGGGCACCGGACGGCCAGGTGTTCTACAAGTATGGGAAAC	6229
QY	6084	TGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTCAAGTTTCAACCTATGACCCAGA	6143
Db	6230	TCTCCAGTTATCAGAGATTGTCTACGACAGTACCGCCGTCACCTTCGGGTATGACGAGA	6289
QY	6144	CGGCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGCTTCACTGCACCATCCGCT	6203
Db	6290	CCACTGGTGTCTTGAAGATGGTCAACCTCCAAAGTGGGGCTTCTCTGCACCATCAGT	6349
QY	6204	ACCGTCAGATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGG	6263
Db	6350	ACCGAAGATTGGCCCCCTGGTGGACAAGCAGATCTACAGGTTCTCCGAGGAAGGCATGG	6409
QY	6264	TCAACGCCGTTTTTGACTACAACCT--ATGACAAACAGCTTCCGGGTGACCAGCATGCAGG	6320
Db	6410	TCAATGCCAGGTTTGACTACACCTATCATATGACAAACAGCTTCCGCATCGCAAGCATCAAGC	6469
QY	6321	CTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTCAAGCA	6380
Db	6470	CCGTCATAAGTGAGACTCCCCCTCCCGTTGACCTCTACCGCTATGATGAGATTTCTGGCA	6529
QY	6381	AGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACAGATCATCACCA	6440
Db	6530	AGGTGGAACACTTTTGGTAAGTTTGGAGTCATCTATTATGACATCAACCCAGATCATCACCA	6589
QY	6441	CAGCTGTGATGACCCACACCAAGCATTTTGATGCAATATGGCAGGATGAAGGAAGTGCAGT	6500
Db	6590	CTGCCGTGATGACCTTCAGCAAAACACTTCGACACCCCATGGCGGATCAAGGAGGTCCAGT	6649
QY	6501	ATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGCGAG	6560
Db	6650	ATGAGATGTTCCGGTCCCTCATGTACTGGATGACGGTGCATATGACAGCATGGGCAGGG	6709
QY	6561	TAGTGAAGAAGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGAGT	6620
Db	6710	TGATCAAGAGGAGCTAAACTGGGGCCCTATGCCAATACCAAGGATACACCTATGACT	6769
QY	6621	ATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACA	6680
Db	6770	ACGATGGGACGGGCAGCTCCAGAGCGTGGCCGTCAATGACCGCCCGACCTGGCGCTACA	6829
QY	6681	GCTACGACCTCAATGGGAACCTTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACAC	6740
Db	6830	GCTATGACCTTAATGGGAATCTCCACTTACTGAACCCAGGCACAGTGTGGCCCTCATGC	6889
QY	6741	CACCTACGGTATGACATCCGGCACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGG	6800
Db	6890	CCTTGGCTATGACCTCCGGGATCGGATAAACAGACTCGGGGATGTGCAGTACAAAATTG	6949
QY	6801	ATGAGGATGGCTTCTTGAGGCAGCGGGCGGTGATATCTTTTGAGTACAACCTCAGCTGGCC	6860

Db 6950 ACGACGATGGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAATTCGAAGGGCC 7009  
QY 6861 TGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCC 6920  
Db 7010 TCCTAAACAAGAGCCTACAACAAGGCCAGCGGTGGAGTGTCCAGTACCGCTATGATGGCG 7069  
QY 6921 TGGGGGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCACCTGTCAGTTCCTCTATGCAG 6980  
Db 7070 TAGGACGGCGGGCTTCCTACAAGACCAACCTGGGCCACCACCTGCAGTACTTCTACTCTG 7129  
QY 6981 ACCTGACCAACCCCAACCAAGTCAACCCACCTGTACAACCACTCCAGCTCTGAGATCACCT 7040  
Db 7130 ACCTCCACAACCCGACCGGCATCACCCATGTCTACATCACTCCAACCTCGAGATTACCT 7189  
QY 7041 CCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGATGAGT 7100  
Db 7190 CACTGTACTACGACCTCCAGGGCCACCTCTTTGCCATGGAGCAGCAGTGGGAGGAGT 7249  
QY 7101 TTTACATAGCTTGTGACACACATCGGGACCCCTCTTGCTGTCTTTAGTGGACAGGTTTGA 7160  
Db 7250 ACTATGTTGCCCTCTGATAACACAGGGACTCCTCTGGCTGTGTTTCAGCATCAACGGCCTCA 7309  
QY 7161 TGATCAAGCAAACTCCTGTACACAGCCCTATGGGGAGATCTACATGGATACCAACCCCAACT 7220  
Db 7310 TGATCAACAGCTGCAGTACACGGCCTATGGGAGATTATTATGACTCCAAACCCCGACT 7369  
QY 7221 TTCAGATCATCATAGGCTACCATGGTGGCCCTCTATGATCCACTCAACCAAGCTGTGCCACA 7280  
Db 7370 TCAGATGGTCAATTGGCTTCCATGGGGACTCTATGACCCCTTGACCAAGCTGGTCCACT 7429  
QY 7281 TGGGCGCGGAGATTATGATGTCTGCGCGGACGCTGGACTAGCCAGACCCAGAGCTGT 7340  
Db 7430 TCACTCAGCGTGATTATGATGTCTGGCAGGACGATGGACCTCCCGAGACTATACCATGT 7489  
QY 7341 GGAAGCACCTTAGTAGCAACAACGTCAATGCTTTTAACTCTATATGTTCAAAAAACAACA 7400  
Db 7490 GGAATAAACGT---GGGCAAGGAGCGGCCCTTTAACTGTATATGTTCAAGAGCAACA 7546  
QY 7401 ACCCCATCAGCAACTCCAGGACATCAAGTGTCTCATGACAGATGTTAACAGCTGGCTGC 7460  
Db 7547 ATCCTCTCAGCAGTGAGCTAGATTTGAAGAATCTACGTGACAGATGTGAAGAAGCTGGCTTG 7606  
QY 7461 TCACCTTTGGATTCCAGCTACACAACGTGATCCCTGGTTATCCCAAAACAGACATGGATG 7520  
Db 7607 TGATGTTTGGATTTCAGCTTAGCAACATCATCTCTGGCTTCCCGAGAGCAAAATGTATT 7666  
QY 7521 CCATGGAACCCCTCCTACGAGCTCATCCACACACAGATGAAAACGAGAGTGGGACAACA 7580  
Db 7667 TCGTGCCCTCCTCCCTATGA-----ATTGTCAGAGAGTCAAGCAAGTGAGAATG 7714  
QY 7581 GCAAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCTTTGTCACT 7640  
Db 7715 GACAGCTCATTAACAGGTGTCCAACAGACAACAGAGAGACATAACCGGCTTCATGGCTC 7774  
QY 7641 TAGAACGGTTTGACCAAGCTCTATGGCTCCACAATCACAGCTGCCAGCAGGCTCCAAAGA 7700  
Db 7775 TGGAA-----GGACAGGTCAATTACTAAAAGTCCACGCCAGCATCCGAGAGAAAGCAG 7828  
QY 7701 CCAAGAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGG 7760  
Db 7829 GTCAGTGGTTTGCCACCAACCGCCCATCATTTGGCAAGGCATCATGTTTGCCATCAAAG 7888  
QY 7761 ATGGCCGAGTGACCACAGACATCATCAGTGTGGCCAAATGAGGATGGCGAAGGTTGCTG 7820  
Db 7889 AAGGGCGGGTGACCAACGGCGGTGTCCAGCATCGCCAGCGAAGATAGCCGCAAGGTGGCAT 7948  
QY 7821 CCATCTTGAACCATGCCCACTACCTAGAGAACTGCATTCACCATGATGGGTGGATA 7880  
Db 7949 CTGTGCTGAACAACGCTACTACCTGGACAAGATGCATACAGCATCGAGGGCAAGGACA 8008  
QY 7881 CCCATTACTTTGTGAACCAAGGACCTTCAGAAAGGTGACCTGGCCATCCTGGGCCTCAGTG 7940

Db 8009 CCCACTACTTTTGTGAAGATTGGCTCAGCCGATGGCGACCTGGTCACACTAGCACCAACCA 8068  
QY 7941 GGGGGCGCGAACCCCTGGAGAATGGGTCAACGTCACTGTGTCCCAGATCAACACAGTAC 8090  
Db 8069 TCGCGCGCAAGGTGCTAGAGAGCGGGGTGAACGTGACCGTGTCCCAGCCACGCTGCTGG 8128  
QY 8001 TTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAGTACGGGGCACTGTGCTTGA 8060  
Db 8129 TCAACGGCAGGACTCGAAGGTTCAAGAACATTGATTCAGTACTCCACGCTGCTGCTCA 8188  
QY 8061 ACACACGCTACGG-----GACAACTTTGGATGAGGAGAGGACACGGGTCCTGGAGC 8111  
Db 8189 GCATCCGCTATGGCCTCACCCCTCCGACACCTGGAGCAAGAGAGGCGCCGCTCCTGGACC 8248  
QY 8112 TGGCCCGCAGAGAGCGCTGCGCCAAAGCGTGGGCCCGCAGCAGCAGAGTGCCTGGGAAG 8171  
Db 8249 AGGCGAGACAGAGGGCCCTGGCAGCGGCTGGGCAAGGAGCAGCAGAAAGCCAGGGACG 8308  
QY 8172 GGGAGGAAGCCCTGCGGCTGGACAGAGGGGGAGAGAGCAGCAGGTGCTGACACAGGGC 8231  
Db 8309 GGAGAGAGGGGAGCGCCTGTGGACTGAGGGCGAGAGAGCAGCAGCTTCTGAGCACCGGC 8368  
QY 8232 GGTCGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCCAGAACTGTGAG 8291  
Db 8369 GCCTGCAAGGGTACGAGGATATTACGTGCTTCCCGTGGAGCAATACCCAGAGCTTGCAG 8428  
QY 8292 ACAGCGCCCAACAACATCCACTTCATGACACAGAGCGAGATGGCGGAGGTGACAGA 8348  
Db 8429 ACAGTAGCAGCAACATCCAGTTTAAAGACAGATGAGATGGAAAGAGTAAACAA 8485

RESULT 9

ADB32023

ID ADB32023 standard; cDNA; 9826 BP.

XX ADB32023;

DT 04-DEC-2003 (first entry)

XX Human FCTR3b cDNA.

Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;  
myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;  
ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;  
renal cell carcinoma; melanoma; clear cell carcinoma;  
granular cell carcinoma; neurological disorder;  
neurodegenerative disorder; nerve trauma;  
familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;  
Gardner syndrome; mental health condition; immunological disorder;  
allergy; asthma; lung disease; reproductive disorder; deafness;  
glycoprotein deficiency; desmoid tumour; turcot syndrome;  
liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;  
diabetes; schistosoma mansoni infection; spinocerebellar ataxia;  
plasmodium falciparum infection; Groenouw's corneal dystrophy;  
lattice corneal dystrophy.

XX Homo sapiens.

PN US2003087816-A1.

XX 08-MAY-2003.

PF 05-MAR-2001; 2001US-00800198.

XX 03-MAR-2000; 2000US-0186592P.

XX (VERM/) VERMET C.

PA (FERN/) FERNANDES E.

PA (SHIM/) SHIMKETS R.

PA (HERR/) HERRMANN J.

PA (MAJU/) MAJUMDER K.

PA (MACD/) MACDOUGALL J.

PA (MISH/) MISHRA V.





QY	3807	GAATGTCA	CCAA	CATCCT	TAGAGCT	GAGGAATA	AAAGATT	TCAGACAT	AGTTCACAGTCCAG	3866
Db	3953	GAATGTG	ACCAGCA	TCTTGG	AGTTAC	GAAATA	AAAGAGTT	TAAACAT	AGCAACACCCAG	4012
QY	3867	CACACAA	ATACTAC	CTGGCC	CACAGAC	CCCATG	AGTGGG	CCCGTCTT	CCCTTCTGACAGCA	3926
Db	4013	CACACAA	GTACTACT	TGGCAG	TGGAC	CCCCGT	GTCCGG	CTCGCTCT	ACGTGTCGACACCA	4072
QY	3927	ACAGCCG	GGGTCTTT	AAAA	TCAAGT	CCACTG	TGTTGGT	TGGTGA	AGGACCTTGTCAAGAACT	3986
Db	4073	ACAGCAG	GGAATCTA	CCGCGT	CAAGTCT	CTGAGT	GGAA	CCAA	AGACCTGGCTGGGAATT	4132
QY	3987	CTGAGGT	GGTTG	CGGGGA	CAGGTG	AC	CAGTGC	CTCCCTTT	TGATGACACTCGCTGCGGGG	4046
Db	4133	CGAAAGT	TGTGG	CAGGGA	CGGAG	AGCAGT	GTCTA	CCCTTT	TGATGAAGCCGCTGCGGGG	4192
QY	4047	ATGTTGG	AAGGCCA	CAGAA	GCCACACT	CA	CCAATCC	CAGGGGTATT	TACAGTGGACAAAT	4106
Db	4193	ATGAGGA	AGGCAT	AGATG	CNA	CCCTGAT	GAGCCG	AGAGGTTATT	TGCAGTAGACAAGA	4252
QY	4107	TTGGGCT	GATCTACT	TTCGT	GGATGG	CACCA	TGATC	GACG	CATCGATCAGAAATGGGATCA	4166
Db	4253	ATGGGCT	CATGTACT	TGTG	CGATGCC	ACCATG	ATCGGA	AGGTTG	ACCAAGATGGAATCA	4312
QY	4167	TCTCCAC	CCCTGCT	CGGCTCTA	ATGATCT	CACATC	AGCCG	CCCACTCAG	CTGATGTTCTG	4226
Db	4313	TCTCCAC	CCCTGCT	CGGCTCTA	ATGATCT	CACATC	AGCCG	CCCACTCAG	CTGATGTTCTG	4372
QY	4227	TCATGG	ATAATTTCC	AGGTAA	AGACTGG	AGTGGT	CTCTG	CAATCT	CTGAAAA	4286
Db	4373	GCATGG	ATGTAG	CCAGGTT	CGTCTG	GAGTGG	CCCAAC	AGACCTT	GTGTCAATCCCATGG	4432
QY	4287	ACAAC	TCACTTTAT	GTCTCG	ACAACA	ATGTG	GTCTCTG	CAATCT	CTGAAAA	4346
Db	4433	ATAAC	TCTTGTAT	GTCTAG	AGAACAA	TGTGAT	CTCTT	CGAATC	ACCGAGAACCAACG	4492
QY	4347	TGGGCAT	TGTG	CGCGGAG	GGCCCATG	CACTG	CCAGGTCC	CTGGCAT	TGACCACTTCCTGC	4406
Db	4493	TCAGCAT	CAATTG	CGGAG	CGCCCCAT	GCACTG	CCAAGTTC	CTGGCAT	TGACTACT---	4549
QY	4407	TAAGCA	AGGTGG	CCATCC	ACGCAAC	CCCTGG	AGTCA	GCAC	CCCGCTTTGGCTGTTT	4466
Db	4550	TCAGCAA	ACTAG	CCATTA	CTCTG	CCCTGG	AGTCA	GCAC	CCCGCTTTGGCTGTTT	4609
QY	4467	ATGGG	TCTGTAT	ATTGCT	GAGACTG	ATG	AGAAAA	AGATCA	ACCGCATCAGCAGGTCA	4526
Db	4610	CTGGG	TCTCTAC	ATCACT	CTG	CCCTT	AGTCA	GCAC	CCCGCTTTGGCTGTTT	

QY	4887	TCACACTCATCAGACAAACAATGGCAACATGGTAAATGTCCGCCGAGACTTACTGGGA	4944
Db	5030	TCACTGAATTGATTGACAATAATGGGAATTCCTGAAGATCCGTGGGACAGCAGTGGCA	5089
QY	4947	TGCCCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGACCAACA	5006
Db	5090	TGCCCCCGTCACCTGCTCATGCCTGACCAACCAGATCATCACCTTCACCGTGGGCACCAATG	5149
QY	5007	GTGCACTCAAGAGTGTGACCAACACAAGGACACAGAGTGGCCATGATGACATACCATGGCA	5066
Db	5150	GAGGCCCTCAAAGTCGTGCCACACAGAACCTGGAGCTTGGTCTCATGACCTATGATGGCA	5209
QY	5067	ATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACAAACATTTTATGAGTACG	5126
Db	5210	AACTGGGCTCCTGGCCACCAAGAGCGATGAAACAGGATGGACGACTTTCTATGACTATG	5269
QY	5127	ACAGCTTTGGCCGCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAA	5186
Db	5270	ACCACGAAGGCCGCTGACCAACGTGACGCCGCCACGGGGTGGTAACCAAGTCTGCACC	5329
QY	5187	GTGATACAGACAGTTCAGTGCATGTCCAGGTAGAGACCTCCA--GCAAGGATGATGTCA	5243
Db	5330	GGGAAATGGAGAAATCTATTACCATTTGACATTTAGAACTCCAAACCGTGATGATGACGTCA	5389
QY	5244	CCATAACCAACCACTGTCTGCCTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCC	5303
Db	5390	CTGTCAACCAACCTCTCTTCAGTAGAGGCCCTCTACACAGTGTACAGATCAAAGTTC	5449
QY	5304	GGAAACAGCTACTACATCGGGCCGATGGCTCCTTSCGGCTGCTGCTGGCCAACGGCATGG	5363
Db	5450	GGAAACAGCTACCAGCTCTGTAAATAATGTTACCTGAGGGTGATGTATGCTAATGGGATGG	5509
QY	5364	AGTGGCGCTGCAGACTGAGCCCCCACTTGTGGCTGGCACCGTCAACCCACCGTGGGCA	5423
Db	5510	GTATCAGCTTCCACAGCAGCCCCCATGTCTAGCGGGCACCATCACCCACCATTGGAC	5569
QY	5424	AGAGGAATGCACGCTGCCCATCGACAAACCGCTCAACCTGGTGGAGTGGCGCCAGCCCA	5483
Db	5570	GCTGCAACATCTCCCTGCCCTATGGAGAAATGGCTTAAACTCCATTGAGTGGCGCCTAGAA	5629
QY	5484	AAGAGCAGGCTCGGGCCAGGTCACTGTCTTTGGCGCCGGCTGCGGTGCACAACCGAA	5543
Db	5630	AGGAACAGATTAAAGGCAAAAGTCACCATCTTTGGCAGGAAGCTCCGGGTCCATGGAAGAA	5689
QY	5544	ATCTCTATCTCTGGACTTTGATCGCGTAACACGCAAGAGAAAGATCTATGATGACCACC	5603
Db	5690	ATCTCTTGTCATTGACTATGATCGAAATATTCGGAATTCGGAAGTGAAGATCTATGATGACCACC	5749
QY	5604	GCAAGTTTCAACCTTCGGATTCTGTACGACAGGCGGGCGGCCAGCCTCTGTGCACCCA	5663
Db	5750	GGAAGTTTCAACCTGAGGATCATTTATGACCAAGTGGGCGGCCCTTCTCTGGCTGCCCA	5809
QY	5664	GCAGCAGGCTGAATGGTGTCAACGTGACATCTCCCTCGGGGTTACATTGCTGGCATCC	5723
Db	5810	GCAGCGGCTGGCAGCTGTCAACGTGTCTACTCTTCAATGGGCGCCTTGCTGGGCTTC	5869
QY	5724	AGAGGGGCATCATGTCTGAAAGAAATGGAATACGACAGGCGGGCCGCGCATCACATCCAGGA	5783
Db	5870	AGCGTGGGGCCATGAGCGAGAGGACAGACATCGACAAGCAAGCCGCGCATCGTGTCCCGCA	5929
QY	5784	TCTTCGCTGATGGGAAGACATGGAGCTACACATCTTAGAGAAAGTCCATGGTCTGCTAC	5843
Db	5930	TGTTCTGCTGACGGGAAGTGTGGAGCTACTCTTACCTTGACAAGTCCATGGTCTCTCTGC	5989
QY	5844	TACACAGCCAGGCGAGTATCTTTGAGTTCGACAAGAAATGACCGCCTCTCTTCTGTGA	5903
Db	5990	TTTCAAGCCCAACGTCAGTATATATTTGAGTATGACTCTCTGACCGCCTCTTGGCGTCA	6049
QY	5904	CGATGCCCAACGTTGGCGGCGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAA	5963
Db	6050	CCATGCCACGCTGGCCCGGCAACAGCATGTCCACACACACCTCCATCGGCTACATCCGTA	6109
QY	5964	ACATCTATCAGCCCCCTGAGGGCAATGCCCTCAGTCTACAGGACTTCACTGAGGATGGGC	6023

Db 6110 ATATTACAAACCCGCTGAAGCAATGCTTCGGTCACTTTGACTACAGTGATGACGGCC 6169  
 QY 6024 ACCTCCTTACACACCTTCTACCTGGGCACTGGCCGAGAGGTGATATACAAGTATGGCAAAC 6083  
 Db 6170 GCATCCTGAAGACCTCCTTTTGGGACCGGACCGCAGGTGTTCTACAAGTATGGAAAC 6229  
 QY 6084 TGTCAAAGCTGGCAGAGACGCTCTATGACACCAAGGTGAGTTTCACTATGACGAGA 6143  
 Db 6230 TCTCAAAGTTATCAGAGATTGTCTACGACAGTACCGCGTCACCTTCGGGTATGACGAGA 6289  
 QY 6144 CGGAGGCATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACTGACCACTCCGCT 6203  
 Db 6290 CCACTGGTGTCTTGAAGATGGTCAACCTCCAAAGTGGGGGCTTCTCCTGCACCATCAGGT 6349  
 QY 6204 ACCGTGAGATTGGGCCCTGATGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGG 6263  
 Db 6350 ACCGAAAGATTGGCCCCCTGGTGACAAAGCAGATCAAGGTCTCCGAGGAAGGCATGG 6409  
 QY 6264 TCAACGCCCGTTTGTACTACAACT--ATGACAAACAGCTTCCGGGTGACGAGCATGCAGG 6320  
 Db 6410 TCAATGCCAGGTTTGTACTACACCTATCATGACAAAGCAGATCAAGGTCTCCGAGGAAGGCATGG 6469  
 QY 6321 CTGTGATCAACGAGACCCACTGCCATTTGATCTCTATCGCTATGATGATGATGATGATGATG 6380  
 Db 6470 CCGTCATAAGTGAGACTCCCTCCCGTTGACCTTACCGCTATGATGAGATTCTTGGA 6529  
 QY 6381 AGACAGAGCAGTTTGGGAAGTTTGGTGTCACTTTATGATGATGATGATGATGATGATGATGATG 6440  
 Db 6530 AGGTGGAACACTTTGGTAAGTTTGGAGTCACTATGATGATGATGATGATGATGATGATGATG 6589  
 QY 6441 CAGCTGTATGACCCACCAAGCATTTTGTGATGATGATGATGATGATGATGATGATGATGATG 6500  
 Db 6590 CTGCCGTGATGACCCCTCAGCAACACTTCGACACCCATGGCGGATCAGGAGGTCCAGT 6649  
 QY 6501 ATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATGATGATGATGATGATG 6560  
 Db 6650 ATGAGATGTTCCGGTCCCTCATGTACTGGATGACGGTGCAATATGACAGCATGGGCAGGG 6709  
 QY 6561 TAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCAATACCACTCGCTACTCCTATGAGT 6620  
 Db 6710 TGATCAAGAGGGAGCTAAACTGGGGCCCTATGCCAATACCAAGAGTACACCTATGACT 6769  
 QY 6621 ATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAAGCCACTCTGGCGCTACA 6680  
 Db 6770 ACGATGGGACGGGCAGCTCCAGAGCGTGGCCGTCATGACCGCCAGCCGCTGCGCTATGC 6829  
 QY 6681 GCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGCTCACAC 6740  
 Db 6830 GCTATGACCTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGCGCCTCATGC 6889  
 QY 6741 CACTACGGTATGACATCCCGCAGCCGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6800  
 Db 6890 CCTTGGCTATGACCTCCGGGATCGGATACCAAGCTCGGGGATGCGAGTGTGAGTGTGAGTGTGAGTGTG 6949  
 QY 6801 ATGAGGATGGCTTCTGAGGCAGCGGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6860  
 Db 6950 ACGACGATGGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAATTCGAAGGGCC 7009  
 QY 6861 TGCTCATCAAGGCCTACAACCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 6920  
 Db 7010 TCCTAACAAAGAGCCTACAACAGCCAGCGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 7069  
 QY 6921 TGGGGCGGCGGTGTCAGCAAGAGAGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 6980  
 Db 7070 TAGGACGGCGGCTTCTTACAAGACCAACCTGGGCCACCACTGCGACTACTTCTACTCTG 7129  
 QY 6981 ACCTGACCAACCCCAAGGTCAACCCACTGTACAACCACTCCAGCTCTGAGATCACCT 7040  
 Db 7130 ACCTCCACAACCCGACGGCATCACCCATGTACAATCACTCCACTCGGAGATTACCT 7189  
 QY 7041 CCCTCTACTACGACTTGAAGGACACCTCTTGGCCATGGAGCTGAGCAGTGGTATGAGT 7100

Db 7190 CACTGTACTACGACCTCCAGGGCCACCTCTTTGCCATGGAGAGCAGCAGTGGGAGGAGT 7249  
 QY 7101 TTATACATAGCTTGTGACAAACATCGGGACCCCTCTTGTCTGTCTTTAGTGGAAACAGGTTTGA 7160  
 Db 7250 ACTATGTTGCCTCTGATAAACACAGGGACTCCTCTGGCTGTGTTTTCAGCATCAACGGCCTCA 7309  
 QY 7161 TGATCAAGCAAATCCTGTACACAGCCCTATGGGAGATCTACATGGATACCAACCCCAACT 7220  
 Db 7310 TGATCAAAACAGCTGCAGTACACGGCCTATGGGAGATTATATGACTCCAACCCCGACT 7369  
 QY 7221 TTGATCATCATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACA 7280  
 Db 7370 TCCAGATGTTCAATTGGCTTCCATGGGACTCTATGACCCCTGACCAAGTGGTCCACT 7429  
 QY 7281 TGGCCCGGCGAGATTATGATGTGCTGGCCGACCGCTGGACTAGCCCGACACGAGCTGT 7340  
 Db 7430 TCACTCAGCGTGATTATGATGTGCTGGCAGGACGATGGACCTCCCGAGACTATACCATGT 7489  
 QY 7341 GGAAGCACCTTAGTAGCAGCAACGTCTATGCTCTTAACTCTATATGTTCAAAAACAACA 7400  
 Db 7490 GGAATAACGT--GGGCAAGGAGCGGCCCTTTAACTGTATATGTTCAAGAGCAACA 7546  
 QY 7401 ACCCATCAGCAACTCCAGGACATCAAGTGTCTTATGACAGATGTTAAACAGCTGGCTGC 7460  
 Db 7547 ATCCTCTCAGCAGTAGTAGATTGTAAGAACTACGTGACAGATGTGAAAAGCTGGCTTG 7606  
 QY 7461 TCACCTTTGGATCCAGCTACAAACGTGATGCTCTGTTATCCCAACCCAGACATGGATG 7520  
 Db 7607 TGATGTTTGGATTTAGCAACATCACTCTGGCTTCCCGAGAGCCAAATGTATT 7666  
 QY 7521 CCATGGAACCCCTCCTAGGACTCATCCACACAGATGAAAACGAGAGTGGGACAACA 7580  
 Db 7667 TCGTGCCTCCTCCCTATGA-----ATTGTACAGAGTCAAGCAAGTGAGAATG 7714  
 QY 7581 GCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCTTTGTACCT 7640  
 Db 7715 GACAGCTCATTTACAGGTGTCCAACAGACAAACAGAGAGACATAACCCAGGCTTTCATGGCTC 7774  
 QY 7641 TAGAACGGTTTGACCAAGCTCTATGGCTCCCAATCACCAGCTGCCAGAGCTCCAAAGA 7700  
 Db 7775 TGGAA-----GGACAGGTCACTACTAAAGCTCCACGCCAGCATCCGAGAGAAAGCAG 7828  
 QY 7701 CCAAGAAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTCCCTTGAAG 7760  
 Db 7829 GTCAGTGGTTTGGCCACCAACCGCCCATCATTTGGCAAGGCATCATGTTTCCATCAAAG 7888  
 QY 7761 ATGCCCGAGTGACCAAGACATCATCAGTGTGGCCAATGAGGATGGGCAAGGGTTGCTG 7820  
 Db 7889 AAGGCGGGTGACCAAGGCGTGTCCAGCATCGCCAGAGATAGCCGCAAGGTGGCAT 7948  
 QY 7821 CCATCTTGAACCATGCCACTACCTAGAGAACCTGCATTCACCATTTGATGGGTGGATA 7880  
 Db 7949 CTGTGCTGAACAACCGCTACTACCTGGACAGATGCACTACAGCATCGAGGGCAAGGACA 8008  
 QY 7881 CCCATTACTTTGTGAAAACCAAGGACCTTCAAGAGGTGACCTGGCCATCCTGGGCTCAGTG 7940  
 Db 8009 CCCACTACTTTGTGAAAGATTGGCTCAGCCGATGGCGACCTGGTCACTACGACCAACA 8068  
 QY 7941 GGGGGCGGCGAACCTTGAGAAATGGGGTCAACGTCACTGTGTCCAGATCAACACAGTAC 8000  
 Db 8069 TCGGCGCGCAAGGTGCTAGAGAGCGGGGTGAACGTGACCGTGTCCAGCCCAACGCTGTGG 8128  
 QY 8001 TTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCAGTACGGGGCACTGTGCTTGA 8060  
 Db 8129 TCAACGGCAGGACTCGAAGGTTTCAAGAAATGAGTTCAGTACTTCCAGCTGTCTGTCTCA 8188  
 QY 8061 ACACACGCTACGG-----GACAACTGTGGATGAGGAGAAAGCACGGTCTCTGGAGC 8111  
 Db 8189 GCATCCGCTATGGCTCAACCCCGACACCTTGACGAAGAGAGAGGCGCCGCTCTGGACC 8248  
 QY 8112 TGGCCCGGCGAGAGCCGTGCGCCAAAGCTGGGCGCCCGGAGCGAGCAGAGACTGCGGGAAG 8171  
 Db 8249 AGCGGAGACAGAGGSCCTTGGGCACGGCCTGGCCCAAGGAGCGAGCAGAGAAAGCCAGGAGC 8308







Db 5053 GAGGCTCAAAGTCGTGTCCACACAGAAACCTGGAGCTTGGTCTCATGACCTATGATGGCA 5112  
 Qy 5067 ATTCCGGCCTTCTGGCAACCAAAAGCAATGAAACCGATGGACAAACATTTTATGAGTACG 5126  
 Db 5113 AACTGGGCTCCTGGCCACCAAGACGATGAAACAGATGGACGACTTTCTATGACTATG 5172  
 Qy 5127 ACAGTTTGGCCCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAA 5186  
 Db 5173 ACCAGAAAGCCCGCTGACCAACGTGACGCGCCACGGGGTGGTAACAGTCTGCACC 5232  
 Qy 5187 GTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTCCA--GCAAGGATGATGCA 5243  
 Db 5233 GGGAAATGGAGAAATCTATTACCATTGACATTGAGAACTCCAACCGTGTGATGACGTCA 5292  
 Qy 5244 CCATAACCAACCACTGTCTGCCTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCC 5303  
 Db 5293 CTGTCAATCACCACCTCTCTCAGTAGAGGCTCTCTACACAGTGGTACACAGATCAAGTTC 5352  
 Qy 5304 GGAACAGCTACTACATCGGGCCGATGGCTCCTTGGGCTGCTGCTGGCCAAACGGCATGG 5363  
 Db 5353 GGAACAGCTACCAGCTCTGTAAATAATGTTACCTCAGGGTGTATGATGCTAATGGATGG 5412  
 Qy 5364 AGTGGCGCTGCAGACTGACGCCCACTTGTCTGGCTGGCACCGTCAACCCACCGTGGCA 5423  
 Db 5413 GTATCAGCTTCCACAGCGAGCCCATGTCTTAGGGGACCATCACCCCCACCATTTGGAC 5472  
 Qy 5424 AGAGGAATGTACGCTGCCCCCACTGCTGGCTGGCACCGTCAACCCACCGTGGCA 5483  
 Db 5473 GCTGCAACATCTCCCTGCTATGGAGAAATGGCTTAAACTCCATTGAGTGGCGCTAAGAA 5532  
 Qy 5484 AAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCGCCGGCTGCGGGTGCACAAACCGAA 5543  
 Db 5533 AGGAACAGATTAAAGGCAAGTCAACCATCTTTGGCAGGAAGCTCCGGGTCCATGGAAGAA 5592  
 Qy 5544 ATCTCCTATCTCTGGACTTGTATCGCGTAAACCGCACAGAGAAGATCTATGATGACCACC 5603  
 Db 5593 ATCTCTTGTCCATTGACTATGATCGAAATATTCGGACTGAAAGATCTATGATGACCACC 5652  
 Qy 5604 GCAAGTTCACCCCTTCGATCTGTACGACAGCGGGGGCGCCAGCCTCTGGTCAACCCA 5663  
 Db 5653 GGAAGTTCACCCCTGAGGATCATTTATGACCAAGTGGCGCCGCCCTTCTCTGGCTGCCCA 5712  
 Qy 5664 GCAGCAGGCTGAATGGTGTCAACGTGACATATCTCCCTGGGGTTACATGCTGGCATCC 5723  
 Db 5713 GCAGCGGCTGGCAGCTGTCAACGTGTCACTACTTCAATGGCGCGCTGGCTGGCTTC 5772  
 Qy 5724 AGAGGGGCATCATGTCTGAAAGAAATGGAATAGCACAGCGGGCGCCGATCACATCCAGGA 5783  
 Db 5773 AGCGTGGGCCATGAGCGGAGGACAGACATCGACAAGCAAGCGCGCATCGTGTCCCGCA 5832  
 Qy 5784 TCTTCGCTGATGGAGACATGGAGCTACACATACTTAGAGAAGTCCATGCTGCTGCTAC 5843  
 Db 5833 TGTTCGCTGACGGGAAAGTGTGGAGCTACTCTACCTTGACAAGTCCATGCTCTCTCTGC 5892  
 Qy 5844 TACACAGCCAGAGGCACTATCTTTGAGTTCGACAAGAAATGACCGCTCTCTTCTGTGA 5903  
 Db 5893 TTCAGAGCCAAACGTGATATATTTGAGTATGACTCTCTGACCGCTCTTGGCGTCA 5952  
 Qy 5904 CGATGCCCAACGTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAA 5963  
 Db 5953 CCATGCCAGCGTGGCCCGGACAGCATGTCCACACACACCTCCATCGGCTACATCCGTA 6012  
 Qy 5964 ACATCTATCAGCCCTGAGGGCAATGCCTCAGTACAGGACTTCACTGAGGATGGGC 6023  
 Db 6013 ATATTTACAACCCCGCTGAAAGCAATGCTTCGGTCACTTTTGACTACAGTGTGACGGCC 6072  
 Qy 6024 ACCTCCTTACACCTTCTACCTGGGCACTGGCGGAGGTGATATACAAGTATGGCAAC 6083  
 Db 6073 GCATCCTGAAGACCTCTTTTGGGCAACCGGACCGCAGGTGTTCTACAAGTATGGGAAAC 6132  
 Qy 6084 TGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCACTTTCACCTATGACGAGA 6143  
 Db 6133 TCTCCAAGTTATCAGAGATTGTCTACGACAGTACCGCCGTACCTTCGGGTATGACGAGA 6192

Qy 6144 CGCAGGCGATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCCACCTGCACCATCCGCT 6203  
 Db 6193 CCACTGGTGTCTTGAAGATGGTCAACCTCCAAAGTGGGGCTTCTCCTGCACCATCAGGT 6252  
 Qy 6204 ACCGTCAGATTGGGCCCCGTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGCATGG 6263  
 Db 6253 ACCGGAAGATTGGCCCCCTGGTGGACAAGCAGATCTACAGGTTCTCCGAGGAAGCATGG 6312  
 Qy 6264 TCAACGCCCGCTTTTGTACTACAACT--ATGACAAACAGCTTCCGGGTGACAGCATGCAGG 6320  
 Db 6313 TCAATGCCAGGTTTGTACTACACTATCATGACAAACAGCTTCCGCTATCGAAGCATCAAGC 6372  
 Qy 6321 CTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTAGGCA 6380  
 Db 6373 CCGTCATAAGTGAGACTCCCTCCCTTCCGCTTACCTTACCGCTATGATGAGATTTCTGGCA 6432  
 Qy 6381 AGACAGAGCAGTTTGGAAAGTTTGGTGTCACTTATGATTAATGACATTAACAGATCATCACA 6440  
 Db 6433 AGGTGGAACACTTTGGTAAAGTTTGGAGTCACTATATGACATCAACAGATCATCACA 6492  
 Qy 6441 CAGCTGTATGACCCCAACCAAGCATTTTGTGATGACCGTCCAGTATGATAACATGGGCGAG 6500  
 Db 6493 CTGCCGTGATGACCCCTCAGCAAAACACTTCGACACCCATGGGCGGATCAAGGAGGTCCAGT 6552  
 Qy 6501 ATGAGATCTTCCGCTCGCTCATGTACTGTGATGACCGTCCAGTATGATAACATGGGCGAG 6560  
 Db 6553 ATGAGATGTTCCGCTCCCTCATGTACTGTGATGACCGTGCATATGACAGCATGGGCGAGG 6612  
 Qy 6561 TAGTGAAGAGGAGCTGAAGGTAGGACCTTACCCCAATACCACTCGCTACTCCTATGAGT 6620  
 Db 6613 TGATCAAGAGGAGCTTAAAACTGGGGCCCTATGCCAATACCAAGTACACCTATGACT 6672  
 Qy 6621 ATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTTGGCGCTACA 6680  
 Db 6673 ACATGGGACGGGAGCTCCAGAGCGTGGCGCTCAATGACCCCGGCTGGCGCTACA 6732  
 Qy 6681 GCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACAC 6740  
 Db 6733 GCTATGACCTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGCGCTCATGC 6792  
 Qy 6741 CACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGTGGTGCAGTGCATACAAGATGG 6800  
 Db 6793 CTTTGGCTATGACCTCCGGGATCGGATACCAAGACTCGGGGATGTCAGTACAAAATTG 6852  
 Qy 6801 ATGAGGATGGCTTCTGAGGACGCGGGGGGTGATATCTTTGAGTACAACTCAGCTGGCC 6860  
 Db 6853 ACACGATGGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAAATCCAAGGGCC 6912  
 Qy 6861 TGCTCATCAAGGCTACAACCGGCTGGAGCTGGAGTGTCAAGTACCGCTACGATGGCC 6920  
 Db 6913 TCCTAACAGAGCGCTACAACAGGCCAGCGGGTGGAGTGTCCAGTACCGCTATGATGGCG 6972  
 Qy 6921 TGGGGCGCGCTGTCCAGCAAGAGCAGCCACAGCCACCTGCTGAGTTCCTTCTATGAG 6980  
 Db 6973 TAGGACGGCGGCTTCTTACAAAGACCAACCTGGGCGCCACCACTGCTGAGTACTTCTACTCTG 7032  
 Qy 6981 ACCTGACCAACCCCAAGGTCAACCACTGTACAAACCACTCCAGCTCTGAGATCACCT 7040  
 Db 7033 ACCTCCACAACCCGACGCGCATCACCCATGTCTACAATCACTCCAATCGGAGATTACCT 7092  
 Qy 7041 CCCTCTACTACGACTTGAAGGACACCTCTTGGCATGGAGTGAAGTGGTGTGATGAGT 7100  
 Db 7093 CACTGTACTACGACTCCAGGGCCACCTCTTGGCATGGAGAGCAGTGGGAGGAGT 7152  
 Qy 7101 TTTACATAGCTTGTGACAACATCGGGACCCCTCTTGTCTGTCTTGTAGTGAACAGGTTTGA 7160  
 Db 7153 ACTATGTTGCCCTGTGATAACACAGGAGTCTCTTGGCTGTGTTGAGCATCAACGGCCTCA 7212  
 Qy 7161 TGATCAAGCAAAATCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCACT 7220  
 Db 7213 TGATCAACAGCTGCAGTACACGGCCTATGGGAGATTTTATGACTCCAACCCCGACT 7272

QY	7221	TT	CAGATCATATAGGCTACCATGGTGGCTCTATATGATCCACTCACCAAGCTTGTCCACA	7281
DB	7273	TCC	AGATGGTTCATGGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGGTCCACT	7333
QY	7281	TGG	CCGGCAGATTATGATGTGCTGGCCGAGCTGGACTAGCCCGAGACCAAGCTGT	7340
DB	7333	TCA	CTCAGCGTGATTATGATGTGCTGGCAGAGCATGGACCTCCCGAGACTATACCATGT	7392
QY	7341	GGA	AGCACCTTAGTAGCAGCAACGTCATGCTCTTTAATCTCTATATGTTCAAAAACAACA	7400
DB	7393	GGAAAA	ACGT---GGGCAAGGAGCGCGCCCTTTAACTGTATATGTTCAAGAGCAACA	7449
QY	7401	ACCC	ATCAGCAACTCCCAGGACATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGC	7460
DB	7450	ATCCT	CTCAGCAGTAGCTAGATTTGAAGAACTACGTGACAGATGTGAAGAGCTGGCTTG	7509
QY	7461	TCAC	CTTTGGATTCCAGCTACACAACGTGATCCCTGGTTATCCCAAAACAGACATGGATG	7520
DB	7510	TGAT	GTTTGGATTTCAGCTTAGCAACATCAATTCCTGGCTTCCCGAGAGCCAAAATGTATT	7569
QY	7521	CCAT	GGAACCTCTCCTACGAGCTCATCCACACACAGATGAAAAACGAGAGTGGGACAACA	7580
DB	7570	TCGT	GCCTCCTCCTATGA-----ATTGTCAGAGAGTCAAGCAAGTCAGAAATG	7617
QY	7581	GCAAG	TCTATCCTCGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCCTTTGTACCT	7640
DB	7618	GACAG	CTCAATTACAGGTGTCCAACAGACAACAGAGAGACATAACCGGCCCTTCATGGCTC	7677
QY	7641	TAGA	ACGGTTTGACCAAGCTCTATGGCTCCACAATCACCAAGCTGCCAGCAGGCTCCAAAGA	7700
DB	7678	TGGAA	-----GGACAGGTCAATTACTAAAAAGCTCCACGCCAGCATCCGAGAGAAAGCAG	7731
QY	7701	CCAAGA	AGTTTGTCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCTTGAAGG	7760
DB	7732	GTCA	CTGGTTTGCCACCACCAACCGCCCATCAATTGGCAAGGCATCATGTTTGCCATCAAAG	7791
QY	7761	ATGG	CCGAGTAGCACACAGACATCATCAGTGTGGCCAAATGAGGATGGCGAAGGTTGCTG	7820
DB	7792	AAGG	CGGGTAGCACACGGGCGTGTCCAGCATCGCCAGCGAAAGATAGCCGAAGGTGGCAT	7851
QY	7821	CCAT	CTTTGAACCATGCCCCACTACCTAGAGAACCTGCACTTCAACATTGATGGGTGGATA	7880
DB	7852	CTGT	GCTGAACAACGCCTACTACCTGGACAAGATGCACATACAGCATCGAGGGCAAGGACA	7911
QY	7881	CCCA	TACTTTTGAAACCAACGAGACCTTCAGAAAGTGAACCTGGCCATCCTGGGCCCTCAGTG	7940
DB	7912	CCCA	CTACTTTGTGAAGATTGGCTCAGCCGATGGCGACCTGGTCACACTAGGCACCACCA	7971
QY	7941	GGGG	CGGGCAACCTTGAGAAATGGGGTCAACGTCACTGTGTCCCAGATCAACACAGTAC	8000
DB	7972	TCGG	CCGCAAGGTGCTAGAGAGCGGGGTGAACGTGACCGTGTCCCAGGCCACGCTGCTGG	8031
QY	8001	TTAA	TGGCAGGACTAGACGCTACACAGACATCCAGTCCAGTACGGGGCACTGTGCTTGA	8060
DB	8032	TCAA	CGGCAGGACTCGAAGGTTTCAAGAACTTGAGTTCAGTACTCCACCGCTGCTGCTCA	8091
QY	8061	ACAC	ACGCTACGG-----GACAACGTTGGATCAGGAGAAAGGCACGGGTCTGAGGAGC	8111
DB	8092	GCAT	CCGCTATGGCCTCACCCCGACACCTGGACGAAGAGAAAGGCCCGCTCCTGGACC	8151
QY	8112	TGGC	CCGGCAGAGAGCGGTGCGCCAAAGCGTGGGCCCGCGAGCAGCAGAGACTGCGGGAAG	8171
DB	8152	AGGC	GAGACAGAGGGCCCTGGGCACGGCCTGGGCCAAGGAGCAGCAGAAAGCCAGGACG	8211
QY	8172	GGGAG	GGAAGGCCTGCGGGCCTGGACAGAGGGGGAGAAAGCAGCAGAGTGTGAGCACAGGGC	8231
DB	8212	GGAG	AGGGGAGCCGCTGTGGACTGAGGGCGAGAGCAGCAGCTTCTTGAGCACCGGGC	8271
QY	8232	GGGT	GCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTGAG	8291
DB	8272	GCGT	GCAAGGGTACGAGGGATATTACGTGCTTCCCGTGGAGCAATACCCAGAGCTTGCAG	8331
QY	8292	ACAG	CGCCAACAACATCCACTTTCATGAGACAGACGAGATGGGCCGAGGTGACAGA	8348

```

||||| 8332 ACAGTAGCAGCAACATCCAGTTTTTAAGACAGAAATGAGATGGGAACAGGTAACAAA 8388
|||||

RESULT 11
ADB32028
ID ADB32028 standard; cDNA; 9729 BP.
XX AC
XX AC
XX AC
DT DT
XX XX
DE Human FCTR3f cDNA.
XX
KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX
OS Homo sapiens.
XX
PN US2003057816-A1.
XX
PD 08-MAY-2003.
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PP 03-MAR-2000; 2000US-0186592P.
XX
PA (VERM/) VERNET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
PA

```

New FCTRX polypeptide and encoding polynucleotide, useful for preventing or treating FCTRX-related disorders, such as cancer, autoimmune, neurodegenerative, gastrointestinal, reproductive and inflammatory diseases.

Claim 9; Page 34-37; 155pp; English.

The invention relates to FCTR<sub>X</sub> polypeptides and the polynucleotides encoding them. The sequences of the invention are useful for the manufacture of a medicament for diagnosing and treating disorders associated with the FCTR<sub>X</sub> polypeptide, such as colorectal cancer, adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain tumours, mammary tumours, human gliomas, astrocytomas, renal cell carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear cell and granular cell carcinomas, neurological disorders, neurodegenerative disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health

CC conditions, immunological disorders, allergy and infection, asthma, lung  
CC diseases, male and female reproductive disorders, deafness, glycoprotein  
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis  
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni  
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,  
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence  
CC represents cDNA encoding an FCTR polypeptide of the invention.  
XX  
SQ

Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Query Match 40.1%; Score 3351.4; DB 9; Length 9729;  
Best Local Similarity 64.6%; Pred. No. 0;  
Matches 5385; Conservative 0; Mismatches 2771; Indels 181; Gaps 19;

QY 35 ATGGACGTGAAGGAGAGAGCCTTACCGCTCGCTGACCCGGCGCGGACGCCGAGCGC 94  
DB 210 ATGGATGTAAGGACCGCGGACACCGCTCTTTGACGAGGACGCTGTGGCAAAGAGTGT 269  
QY 95 CGCTACACGAGCTGTCGCGGACAGCGAGGAGGCAAG---CCCGCAGAAATCGTAC 151  
DB 270 CGCTACACAAGCTCTCTTGACAGTGAGGACTGCCGGTGTGCCACACAGAAATCCTAC 329  
QY 152 AGCTCCAGGAGACCTGTAAGGCTTACGACGAGGACGCCGCTAGCTATGGCAGCCGC 211  
DB 330 AGCTCCAGTGAGACTCTGAAGGCTTATGACATGACAGGAGTGAATGAAACCGA 389  
QY 212 GTCAAGGACATTTGTCGCGGAGGCGGAGGAAATTTGCCGACAGGTGCCAATTTCACC 271  
DB 390 GTCACAGACCTCATCCACGGGAGTCAGATGAGTTTCTTAGACAGGAACCAATTTCACC 449  
QY 272 CTGCGGGAGCTGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCTGTACCGGACAGAC 331  
DB 450 CTTGCCGAACCTGGGCACTGTGA---GCCCTCCCCACACCGAAGGGGCTACTGCTCCGAC 506  
QY 332 ATTGGCCTGCCCAATGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCT 391  
DB 507 ATGGGATCCTTACCAGGGCTACTCCCTTAGCACAGGCTGTGACGCGGACTCCGACACC 566  
QY 392 GACACGGTGTCTCCCTGAGCACCCCGTGGCTCTGTGGGGCCGAGACACCGTCAAGG 451  
DB 567 GAGGAGGGATGTTCCAGAACACGCCATCAGACTGTGGGGCAGGGATAAATCCAGG 626  
QY 452 GCAGCTCCTGCTGTCAGCACCCCGTGGCTCTGTGGGGCCGAGACACCGTCAAGG 511  
DB 627 CGCAGTTCCGGCTGTCCAGTCTGTGAATACTCGGCCCTTACCCCTGACTGACTCTGACAAC 686  
QY 512 GAAACACTGAGACTGATCATCCGGCGGCC---TGCAGAACACCGCGCGCTCCGGAC 567  
DB 687 GAAACAAATCAGATGATGAGAACGGTCTGCCCATTTCCACCTACATCCTCGCTAGTCTC 746  
QY 568 GCCGCGCGCGCTCTCGACGCGCCCAACACCCCAACAGCACGCGGCTCCCAATAA 627  
DB 747 CTCCCATCTGCTCAGCTGCCTAGCTCCCATTAATCCTCCACAGTGTAGTGC-AGATGCC 805  
QY 628 CTCCTGAACCGGGCAACTTCACGCGAGGAGCAACCCAGCGCGGCCCCCAGCGGACCA 687  
DB 806 ATTGCTAGACAGCAACACCTCCCATCAATCATGGACACCAACCTGATGAGGAATTC 865  
QY 688 CTCGCTCTCCGAGAGCCCTCTCGCGCGCGCGCCAGGAGCCTGCGCCACGCGGAGAA 747  
DB 866 CCCCAATTCACTGCTCAGAGCATGTCTCAGGGCCCCAGCAAGCCTCCAGCAGTGGCCC 925  
QY 748 CTGGCTGCTCAACAGCAACATCCCTCTGGAGACCAGGAACCTAGGCAAGCAGCCATTCT 807  
DB 926 TCCGAACCCACAGCCAGTCGACTCTGAGGCC-----CCTCTCCC 967  
QY 808 AGGACATTGGAGGACAACTTCATTGAGATGGACATTTCTCGGCGCTCCCGCCATGATGG 867  
DB 968 ACCCCCTCACACACACGCTGTCTCCATCACCACCTCGTCCGCAACTCCCTCAACAGGAA 1027  
QY 868 GGCTTACAGTGACGGGCACTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGACAC 927  
DB 1028 CTCACCTGACCAATCGCGGAGTCAAGTCCACGCCCCCGGCCCCAGGCCCCCAATGACCTGGC 1087

QY 928 CACATCACCGGGTACCCACTGACGTCCAGACAGTGTACTCTCTCCGCCCCGACCCCT 987  
DB 1088 CACCACACAGAGTCCGTTAGCTTCAGGACAGTGGTGTAAACAGCAACAGTGCCTACT 1147  
QY 988 GCCCGCAGCACCTTCGCCCGCGCGCTTAAACCTCAAGAAGCCCTCAAGTACTGTAA 1047  
DB 1148 GGAGACCGGCACTTCTCTTCAAGACCTCTCTCGGGAGCACACCT-----T 1195  
QY 1048 CTGGAAGTCCGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGC 1107  
DB 1196 GTTCAGCAGCTCTTCCCGGGATACCTTTGACCTCAGGAACGGTTACACGCCCGCC 1255  
QY 1108 ATACTTTGTGGCATGACACCTGTTTGGCTAAACTGGCACCTGACGCCGATGGAGGGCA 1167  
DB 1256 CCGCTGCTGCCAGGAATACTTTCTCCAG-----GAAGCTTTCAA 1297  
QY 1168 GATGTATGAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCC-C 1226  
DB 1298 GCTGAAGAAAGCCCTCCAAATACTGACGTGGAATGTGTGCCCTCTCCGCCATTGCCGC 1357  
QY 1227 TATACCCCTCAGGGGCACTGGCTTAGACACCCCTGACAGGAAGCAAGGAACACACAG 1286  
DB 1358 GGCCCTCCTCTGGCTATTTTGTGGCTATTTTATAGTGCCTGCTGTTGAAAAACAG 1417  
QY 1287 AAGGAAAGCCAGTAGTTTCTTTCCAGAGGACAGTTTTCATAGATTTCTGGAGAAATTTGATG 1346  
DB 1418 CAG-----CATAGACAGTGGTGAAGCAGAAG 1443  
QY 1347 TGGGAAGGCGAGCTCCAGAAAGATTCTCTCTGGCAGCTTTCTGGAGATCTCAAGTGTTC 1406  
DB 1444 TTGGTCGGCGGTAACACAAAGATCCACAGGGGTGTTTGGAGGTCACAAAATTCACA 1503  
QY 1407 TAGACCATCTCTGTCATCTGAAATTCATATGTGTCTCTGGGAAAGCAGCCCTGTTGGCA 1466  
DB 1504 TCAGTCAGCCCGAGTCTTAAAGTTCAACATCTCCCTCGGGAAGGACGCTCTCTTTGGTG 1563  
QY 1467 TTTATGGCAGAAAGGCTCCCTCTTCACATACACAGTTTGTGAGTGTGGAGTGTGG 1526  
DB 1564 TTTACATAAGAAAGAGGACTTCCACCATCTCATGCCAGTATGACTTATGGAACGTCTGG 1623  
QY 1527 ATGGCAGGAGGCTCTAACCCAGGAGCGCGGAGCCTTAGAGGGACCCCGCCAGTCTC 1586  
DB 1624 ACGGA-----AGGAGAGTGGAGTGTGTTGAGTCTCCAGGGAACGCC 1668  
QY 1587 GGGAACTGTGCCCCCTCCAGCCATGAGACAGGCTTTCATCCAGTATTTGGATTTCAGGAA 1646  
DB 1669 GGAGCATACAGACCTTGGTTCAGATGAAGCCGTGTTTGTGAGTACCTGGATGTGGGC 1728  
QY 1647 TCTGGCACTTGGCTTTTACAATGACGGAAGGAGTCAAGAGTGGTTTCTTCTCACCA 1706  
DB 1729 TGTGGCATCTGGCCTTCTACAATGATGGAAGAAAGACAAAGAGATGGTTTCTTCAATCTG 1788  
QY 1707 CTGCCATTGAGTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGAATGCTATCT 1766  
DB 1789 TTGTCTTAGATTAGTGCAGGACTGTCCACGTAACCTGCCATGGGAATGGTGAATGTGTGT 1848  
QY 1767 CTGGGACCTGCCACTGCTTCTCTGGGTTTCTGGCCCCCAGTGTGGCAGAGCCTCTCTGCC 1826  
DB 1849 CCGGGGTGTGTCACTGTTTCCAGGATTTCTAGGAGCAGACTGTGTAAAGTGCCTGCC 1908  
QY 1827 CCGTGTCTGTAGCGGAATGGCCAATACATGAAAGGCGAGATGCTTGTGCCACAGTGGCT 1886  
DB 1909 CTGTCTGTGAGTGGGAATGGACAATATTTCTAAAGGACGTCGTCAGTGTACAGCGGCT 1968  
QY 1887 GGAAGGCGTGAAGTGCATGTGCCACCAACCAAGTGTATCGATGTGGCCTGCAGCAACC 1946  
DB 1969 GGAAGGTCAGAGTGCAGCTGCCCATGAATCAGTGCATCGATCCTTCTCTCGGGGGCC 2028  
QY 1947 ATGGCACCTGTCATCAGGGCACCTGCATCTGCAACCCCTGGTACAAAGGCGAGAGCTGTG 2006  
DB 2029 ACGGCTCCTGCATTGATGGGAATGTGTCTGCTCTGCTGCTACAAAGGCGAGCACTGTG 2088



Db 4216 TCTCCACCCTGCTGGGCTCCAATGACCTCACTGCGCTCCGGCCGCTGAGCTGTGATTCCA 4275  
QY 4227 TCATGGATATTTCCAGGTAAGACTGGAGTGGCCACAGACTTAGCCATCAACCCCAATGG 4286  
Db 4276 GCATGGATGTAGCCCAAGGTTGCTCTGGAGTGGCCAAACAGACCTTGCTGTCAATCCCATGG 4335  
QY 4287 ACAACTCACTTTATGCTCGACAAACAATGTGGTCTCTGCAAAATCTCTGAAAAACCAACAGG 4346  
Db 4336 ATAACTCCTTGATGTCTAGAGAAACAATGTATCCTTCGAATACCGAGAACCAACCAAG 4395  
QY 4347 TGGCCATTGTGCGCGGAGGCCCATGCACTGCCAGGTCCCTGGCATGTGACCACTTCCTGTC 4406  
Db 4396 TCAGCATCATTTGGGACGCCCCCATGCACTGCCAAGTTCCTGGCATGACTACT--CAC 4452  
QY 4407 TAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCCACCGCTTTGGCTGTTTCACACA 4466  
Db 4453 TCAGCAAACTAGCCATTCACTCTGCCCTGGAGTCAGCCAGTGCCTATGCCCATTCTCTACA 4512  
QY 4467 ATGGGGTCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCA 4526  
Db 4513 CTGGGGTCTCTACATCACTGAGACAGATGAGAGAAAGATTAAACGCTCTACGCCAGGTAA 4572  
QY 4527 CCACCTAGTGAGAGATCTCACTCGTTGCTGGGGCCCCCAGTGGTGTGACTGTAAAAATG 4586  
Db 4573 CAACCAACGGGAGATCTGCTCTTTAGCTGGGGCAGCCTCGGACTGCGACTGCAAAAAACG 4632  
QY 4587 ATGCCAACTGTGATTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATA 4646  
Db 4633 ATGTCAATTGCAACTGCTATTTCCAGGAGATGATGCCCTACGCGACTGATGCCATCTTGAATT 4692  
QY 4647 CCCCATTCTCTGCTGCTGTGTGATGGGAGCTCTACGTGGCCGACCTTGGGAACA 4706  
Db 4693 CCCCATTCTCTGCTGCTGTGTGATGGGAGCTCTACGTGGCCGACCTTGGGAACA 4752  
QY 4707 TCCGAATTCCGTTTATCCGGAAGAACCAAGCCTTTCTCTCAACACCCAGAACATGTATGAGC 4766  
Db 4753 TTCGGATCAGGGCGGTGAGCAAGAACCAAGCCTGTTCTTAATGCTTCAACCATGATGAGG 4812  
QY 4767 TGTCTTCAACCAATTGACAGGAGCTCTATCTGTTTGTATACCAACCGGCAAGCCTGTACA 4826  
Db 4813 CTGCATCCCCCGGAGAGCAGGAGTTATATGTTTCAACGCTGATGGCATCCACCAATACA 4872  
QY 4827 CCCAAGCCTGCCACAGGAGACTACCTGTACAACTTCACTACACTGGGAGCGGCGACA 4886  
Db 4873 CTGTGAGCCTGGTGACAGGGGAGTACTTGTACAAATTTACATATAGTACTGACAAATGATG 4932  
QY 4887 TCACACTCATCAGACAAACAATGGCAACATGGTAAATGTCCGCGGAGACTCTACTGGGA 4946  
Db 4933 TCACCTGAATTGATTGACAAATAATGGGAATCCCTGAAGATCCGTCGGGACAGCAGTGGCA 4992  
QY 4947 TGCCCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGCAACAA 5006  
Db 4993 TGCCCCGTCACTGCTCATGCTGACAAACAGATCATCACCTCACCGTGGGCAACCAATG 5052  
QY 5007 GTGCACTCAAGAGTGTGACCAACAAGGACAGAGTTGGCCATGATGACATACCATGGCA 5066  
Db 5053 GAGGCCCTCAAAGTGTGTCCACACAGAACCTGGAGCTTGGTCTCATGACCTATGATGGCA 5112  
QY 5067 ATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAAACGGATGGACAAACATTTATGAGTACG 5126  
Db 5113 ACACCTGGGCTCCTGGCCACCAAGAGCGATGAAACAGGATGGACACTTCTATGACTATG 5172  
QY 5127 ACAGCTTTGGCCGCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAA 5186  
Db 5173 ACCACGAAGGCCGCTGACCAACGTCAGCGGCCCCACGGGGGTGTAACCACTCTGCACC 5232  
QY 5187 GTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTTCCA--GCAAGGATGATGCA 5243  
Db 5233 GGGAAATGGAGAAATCTATTACCATTTGACATTTGAGAACTCCAACCGTGTATGATGACGTC 5292  
QY 5244 CCATAACCAACCACTGTCTGCTCAGGCGCCTTCTACACACTGTCTGCAAGACCAAGTCC 5303  
Db 5293 CTGTCAATCACCAACCTCTCTTTCAGTAGAGGCCTCTACACAGTGGTACAAGATCAAGTTC 5352

QY 5304 GGAACAGCTACTACATCGGGGCCGATGGCTCTCTTGGGCTGCTGCTGGCCAAACGGCATGG 5363  
Db 5353 GGAACAGCTACCAGCTCTGTAAATAATATGGTACCCCTGAGGGTGATGTATGTAATGGATGG 5412  
QY 5364 AGGTGGCGCTGCAGACTGAGCCCCACATTTGCTGGCTGGCACCGTCAACCCCAACCGTGGCA 5423  
Db 5413 GTATCAGCTTCCACAGCGAGCCCCATGTCTAGCGGGCACCATCACCCCAACCATTTGGAC 5472  
QY 5424 AGAGGAATGTACGCTGCCCATCGACAAACGCGCTCAACCTGGTGGAGTGGCGCCAGCGCA 5483  
Db 5473 GCTGCAACATCTCCCTGCTATGGAGAAATGGCTTAAACTCCATTGAGTGGCGCTAAGAA 5532  
QY 5484 AAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGCTGCACAACCGAA 5543  
Db 5533 AGAACAGATTAAAGGCAAAAGTCAACCATCTTTGGCAGGAAGCTCCGGTCCATGGAAGAA 5592  
QY 5544 ATCTCCTATCTCTGGACTTTGATCGCGTTAAACACGACACAGAGAAAGATCTATGATGACCAC 5603  
Db 5593 ATCTCTTGCTCACTGATGATCGAAATATTCGACTGAAAGATCTATGATGACCAC 5652  
QY 5604 GCAAGTTCAACCTTCGGATTCTGTACGACGAGCGGGCGGCGGCTCTGCTGTCACCCA 5663  
Db 5653 GGAAGTTCAACCTTGAGGATCATTTATGACAGGTGGCGGCCCTTCTCTGCTGCCCA 5712  
QY 5664 GCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATCC 5723  
Db 5713 GCAGCGGCTGGCAGCTGTCAACGTGTCTACTTCTCAATGGGCGCTGGTGGCTTC 5772  
QY 5724 AGAGGGGCATCATGTCTGAAAGAAATGGAATACGACAGGGGGCGCATCATCCAGGA 5783  
Db 5773 AGCGTGGGGCATGAGCGAGAGGACAGACATCGAACAGAGCGCGCATCGTGTCCCGCA 5832  
QY 5784 TCTTCGCTGATGGAAAGACATGGAGCTACACATACCTTAGAGAAAGTCCATGGTGTCTAC 5843  
Db 5833 TGTTGCTGACGGGAAAGTGTGGAGTACTCTTACCTTGACAAAGTCCATGGTCTCTCTGC 5892  
QY 5844 TACACAGCCAGAGGCAATATCTTTGAGTTGACAAAGATAGCCGCTCTCTTCTGTGA 5903  
Db 5893 TTCAGAGCCCAACGTCAGTATATATTGAGTATGACTCCTCTGACCGCCTCTTGCCTGCA 5952  
QY 5904 CGATGCCCAACGTCGGCGGACAGACTAGAGACCATCCGCTCAGTGGGCTACTACAGAA 5963  
Db 5953 CCATGCCAGCGTGGCGGACAGCATGTCCACACACACACCTCCATCGGCTACATCCGTA 6012  
QY 5964 ACATCTATCAGCCCCCTGAGGGCAATGCTCAGTCTCAGTACAGGACTTCACTGAGGATGGC 6023  
Db 6013 ATATTACAAACCCGCTGAAAGCAATGCTTCGGTCTCTTTGACTACAGTATGACGGCC 6072  
QY 6024 ACCTCCTTCAACCTTCTACCTGGGCACTGGCGGAGGTGATATACAAAGTATGGCAAAC 6083  
Db 6073 GCATCCTGAAGACCTCTTTTGGGCAACCGGACCGCAGGTGTTCTACAAAGTATGGAAAC 6132  
QY 6084 TGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTGAGTTTCACTTATGACGAGA 6143  
Db 6133 TCTCCAAGTTATCAGAGATTGTCTACGACAGTACCGCGTCACTTCCGGTATGACGAGA 6192  
QY 6144 CGGCAGGCTGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACTGCACCATCCGCT 6203  
Db 6193 CCACTGGTGTCTGAAGATGGTCAACCTCCAAAGTGGGGCTTCTCTGCACCATCAGGT 6252  
QY 6204 ACCGTGAGATTGGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGG 6263  
Db 6253 ACCGGAAGATTGGCCCCCTGGTGGACAAAGCAGATCTACAGGTCTCCGAGGAAGGCATGG 6312  
QY 6264 TCAACGCCGCTTTTGACTACAACT---ATGACAAACAGCTTCCGGGTGACCAAGCATGCAGG 6320  
Db 6313 TCAATGCCAGGTTTGACTACACCTTATCATGACAAACAGCTTCCGATCGCAAGCATCAAGC 6372  
QY 6321 CTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCGTATGATGATGTGTGAGGCA 6380  
Db 6373 CCGTCATAAGTGAGACTCCCTCCCGCTTGACCTCTACCGCTATGATGAGATTTCTTGCA 6432

QY 6381 AGACAGACAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACCAGATCATCACCA 6440  
DB 6433 AGGTGAACACTTTGGTAAGTTTGGAGTCATCTATTATGACATCAACCAGATCATCACCA 6492  
QY 6441 CAGCTGTCAATGACCCACACACCAAGCAATTTTGATGCATATGGCAGGATGAAGGAAGTGCAGT 6500  
DB 6493 CTGCGGTGATGACCTTCAGCAAAACACTTCGACACCCATGGCGGATCAAGGAGTCCAGT 6552  
QY 6501 ATGAGATCTTCCGCTCGCTCATGTACTTGATGACCCGTCAGTATGATAAATCAATGGGCGAG 6560  
DB 6553 ATGAGATGTTCCGGTCCCTCATGTACTTGATGACCGGTGCAATATGACAGCATGGGCGAGG 6612  
QY 6561 TAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGTACTCTCTATGAGT 6620  
DB 6613 TGATCAAGAGGGAGCTAAAACTGGGCCCCCTATGCCAATACCAAGTACACCTATGACT 6672  
QY 6621 ATGATGCTGACGGCCAGCTGACAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACA 6680  
DB 6673 ACGATGGGACGGGCAGCTCCAGAGCGTGGCCGTCAATGACCGCCGACCTGGCGCTACA 6732  
QY 6681 GCTACGACCTCAATGGGAACCTTGCACTTACTGAGCCCTGGGAACAGTGCAACGCTCACAC 6740  
DB 6733 GCTATGACCTTAATGGGAATCTCCACTTACTGAACCCAGCAACAGTGTGCGCCTCATGC 6792  
QY 6741 CACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGG 6800  
DB 6793 CCTTGGCTATGACCTCCGGATCGGATAACAGACTCGGGGATGTGCAGTACAAAATTG 6852  
QY 6801 ATGAGATGGCTTCTCTGAGCAGCGGGGTGGCAGCTGGAGTGCAGGTACCGCTACGCTGGCC 6860  
DB 6853 ACGACGATGGCTATCTGTGCCAGAGAGGGTCTGACATCTTTCGAATACAATTCGAAGGGCC 6912  
QY 6861 TGCTCATCAAGGCCTACAACCGGGGTGGCAGCTGGAGTGCAGGTACCGCTACGATGGCC 6920  
DB 6913 TCCTAACAAGAGCCTACAACAAGCCAGCGGTGGAGTGTCCAGTACCGCTATGATGGCG 6972  
QY 6921 TGGGGCGCGCGTGTCCAGCAAGACAGCAGCCACAGCCACCACTGCAGTCTCTCTATGCAG 6980  
DB 6973 TAGGACGGGGCTTCTCTACAAGACCAACCTGGGCCACCACTGCAGTACTTCTACTCTG 7032  
QY 6981 ACCTGACCAACCCCAACCAAGGTCAACCACTGTATACAACCACTCCAGCTCTGAGATCACCT 7040  
DB 7033 ACCTCCACAACCCGACCGCGCATACCCCATGTCTACAATCACTCCAACCTCGGAGATTACCT 7092  
QY 7041 CCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGTGATGAGT 7100  
DB 7093 CACTGTACTACGACCTCCAGGGCCACCTCTTTGCCATGGAGCAGCAGTGGGGAGGAGT 7152  
QY 7101 TTTACATAGCTTGTGACAAACATCGGACCCCTCTTTGGCTGCTTTAGTGGAAACAGGTTTGA 7160  
DB 7153 ACTATGTGCTCTGTATAACACAGGGACTCCTCTGGCTGTGTTTCAGCATCAACGGCCTCA 7212  
QY 7161 TGATCAAGCAAAATCCTGTACACAGCCTATGGGGAGATCTACATGGATACCAACCCCAACT 7220  
DB 7213 TGATCAAAACAGCTGCAGTACACGGCCTATGGGGAGATTTTATATGACTCCAACCCCGACT 7272  
QY 7221 TTCAGATCATATAGGCTACCATGGTGGCCTCTATGATCCACTCAACCAAGCTTGTCCACA 7280  
DB 7273 TCCAGATGGTCAATGGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGGTCCACT 7332  
QY 7281 TGGGCGGCGAGATATGATGTGCTGGCGGACGCTGGACTAGCCCGACAGACCGAGCTGT 7340  
DB 7333 TCACCTACGCGTATGATGTGCTGGCAGGACGATGGACCTCCCGCAGACTATACCATGT 7392  
QY 7341 GGAAGCACCTTAGTAGCAGCAACGTCATGCTCTTTAATCTCTATATGTTCAAAAACAACA 7400  
DB 7393 GGAAAAACGT---GGCAAGGAGCGCGGCCCTTTAACCTGTATATGTTCAAGAGCAACA 7449  
QY 7401 ACCCCATCAGAACTCCAGGACATCAAGTGTCTTCATGACAGATGTTAACAGCTGGCTGC 7460  
DB 7450 ATCTCTCAGCAGTGAGCTAGATTTGAAGAACTACGTGACAGATGTGAAAGCTGGCTTG 7509  
QY 7461 TCACCTTTGGATTCCAGCTACACAACGATGATCCCTGGTTATCCCAACCAAGACATGGATG 7520

DB 7510 TGATGTTTGGATTTTCAGCTTAGCAACATCATTTCTCTGGCTTCCCGAGAGCCAAAATGTATT 7569  
QY 7521 CCATGGAACCCCTCCTACGAGCTCATCCACACAGATGAAAACGCGAGGATGGGACAACA 7580  
DB 7570 TCGTGCCTCCTCCCTATGA-----ATTGTCAGAGAGTCAAGCAAGTGAGATG 7617  
QY 7581 GCAAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCCCTTTGTCACT 7640  
DB 7618 GACAGCTCATTAACAGGTGTCCAACAGACAACAGAGAGACATAAACAGGCCCTTTCATGGCTC 7677  
QY 7641 TAGAACGGTTTGACCAGCTCTATGGCTCCACAATCAACAGCTGCCAGCAGGCTCCAAAAGA 7700  
DB 7678 TGGAA-----GGACAGGTCAATTACTAAAAAGCTCCACGCCAGCATCCGAGAGAAAAGCAG 7731  
QY 7701 CCAAGAAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGG 7760  
DB 7732 GTCACGTGTTTGCCACCACCCACGCCCATCATTTGGCAAGGATCATGTTTGGCCATCAAAG 7791  
QY 7761 ATGGCCGAGTGACCACAGACATCATCATGAGTGTGGCCAAATGAGATGGCGGAAGGTTGCTG 7820  
DB 7792 AAGGGCGGTGACCACCGGCGTGTCCAGCATCGCCAGCGAAGATAGCCGAAGGTGGCAT 7851  
QY 7821 CCATCTTGAACCATGCCACTACCTAGAGAACTGCACATTACACATTTGATGGGTGGATA 7880  
DB 7852 CTGTGCTGAACAACGCCTACTACCTGGACAAGATGCACATACAGCATCGAGGGCAAGGACA 7911  
QY 7881 CCCATTACTTTGTGAAACCAGGACCTTCAGAAGGTGACCTGGCCATCCTGGCCTCAGTG 7940  
DB 7912 CCCACTACTTTGTGAAGATTGGCTCAGCCGATGGCGACCTGGTCACACTAGGCACCACCA 7971  
QY 7941 GGGGGCGGCGAACCCCTGGAGAAATGGGTCAACGTCACTGTGTCCAGATCAACACAGTAC 8000  
DB 7972 TCGGCCGCAAGGTGCTAGAGAGCGGGGTGAACGTGACCGTGTCCAGCCACCGTCTG 8031  
QY 8001 TTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGA 8060  
DB 8032 TCAACGGCAGGACTCGAAGTTTCAAGAACATTTAGTTCAGTACTCCAGCTGTGCTCA 8091  
QY 8061 ACACACGCTACGG-----GACAACGTTGGATGAGGAGAGGCAAGGCTCCTGGAGC 8111  
DB 8092 GCATCCGCTATGGCCTCACCCCGGACACCTTGGACGAAGAGAGGCGCGCTCCTGGACC 8151  
QY 8112 TGGCCCGGACAGAGCCGTGCGGCCAAGCGTGGGCCCGGAGCAGCAGAGACTGCGGGAAG 8171  
DB 8152 AGCGAGACAGAGGGCCCTTGGGCAACGGCTGGGCCAAGGAGCAGCAAGAACCCAGGGACG 8211  
QY 8172 GGGAGGAAGCCTCGGGCCTGGACAGAGGGGGAGAAAGCAGCAGTGTGCTGAGCACAGGGC 8231  
DB 8212 GGAGAGAGGGAGCCGCTGTGGACTGAGGGCGAGAAAGCAGCAGTCTTGAGCACCCGGC 8271  
QY 8232 GGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTGTCAG 8291  
DB 8272 GCGTGCAAGGCTACGAGGGATATTACGTGCTTCCGTGGAGCAATACCCAGAGCTTGCAG 8331  
QY 8292 ACAGCGCCAACAACATCCACTTTCATGAGACAGAGCGAGATGGGCCGAGGTGACAGA 8348  
DB 8332 ACAGTAGCAGCAACATCCAGTTTAAAGACAGAATGAGATGGGAAAGAGGTAACAAA 8388

RESULT 12

AAK51828

ID AAK51828 standard; cDNA; 13202 BP.

XX AAK51828;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 373.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
XX vaccine; peptide therapy; stem cell growth factor; hematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukemia;  
KW

KW nervous system disorder; arthritis; inflammation; ss.  
XX Homo sapiens.  
OS WO200157190-A2.  
XX PD 09-AUG-2001.  
XX PF 05-FEB-2001; 2001WO-US0004098.  
XX PR 03-FEB-2000; 2000US-00496914.  
PR 27-APR-2000; 2000US-00560875.  
PR 20-JUN-2000; 2000US-00598075.  
PR 19-JUL-2000; 2000US-00620325.  
PR 01-SEP-2000; 2000US-00654936.  
PR 15-SEP-2000; 2000US-00663561.  
PR 20-OCT-2000; 2000US-00693325.  
PR 30-NOV-2000; 2000US-00728422.  
XX (HYSE-) HYSEQ INC.  
XX PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
DR P-PSDB; AAM78695.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX  
PS Claim 1; Page 1414-1426; 6221pp; English.  
XX  
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
XX sequence listing were missing at the time of publication  
SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;  
Query Match 39.2%; Score 3274.2; DB 4; Length 13202;  
Best Local Similarity 65.7%; Pred. No. 0;  
Matches 4910; Conservative 0; Mismatches 2498; Indels 67; Gaps 7;  
QY 880 CGGGCACTTCCTCTCAAGCCTGGA---GGCACCTCCCGCTCTCTGACCATCACC 936  
DB 1164 CAGGCATTTCCCTGTTCAACATGGATCTGGTTCCTCTGCGATCTTCAGTGCAGCCAGTCA 1223  
QY 937 AGGGTACCCACTGACGTCCAGCACAGTACTCTCTCCGCGCCGACCCCTGACCTGGAAGTG 996  
DB 1224 GAACTACCCCTTGACATCCAATACCGTGTACTCGCCCTCCAGGCCTCTTCTCTCGAAG 1283  
QY 997 CACCTTCGCGCGCGCCCTTTAACTCAAGAACCCCTCCAACTGACTGTAACTGGAAGTG 1056  
DB 1284 CACCTTTTCCCGACCTGCCTTTTACCTTTAAACAAACCTTACAGGTGCTGCAACTGGAAGTG 1343  
QY 1057 GCAGCCCTGAGCCCATCGTCACTCAGCCACTCTGGTCACTCTGCTGGCATACTTTGT 1116  
DB 1344 CACAGCATTGAGCGCCACTGCAATCACAGTGAATTTGGCCTTGTACTAGCCTATGTGAT 1403  
QY 1117 GGCCATGCACCTGTTTGGCCTAAACTGGACCTGCGAGCCGATGAGGGGCGAGATGATGA 1176  
DB 1404 TGCAGTGCATTTGTTGCGCCTGACTTGGCAGTTGCAACCACTGTAAGGAGAGCTGTATGC 1463

QY 1177 GATCACGGAGGACACAGCCAGCAGTGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCCTC 1236  
DB 1464 AAATGGAGTTAGCAAGGAAACAGGGGACCGAGTCCATGGACACTACTACTCTCCAAT 1523  
QY 1237 AGGGGGCACTGGCTTAGAGACCCCTGACAGGAAGGCAAGGAACACAGAAAGAAAGCC 1296  
DB 1524 TGGAGGAAAAGTTTCTGATAAATCAGAGAAAAG----- 1558  
QY 1297 CAGTAGTTTCTTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTTGATGTGGAAAGCG 1356  
DB 1559 -----TGTTTCAGAAAGGACGGGCGATAGACACTGGAGAAGTTGATGTGGTGCACA 1610  
QY 1357 AGCCTCCAGAAAGATTCTCTCTGGCACCTTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416  
DB 1611 GGTCAATGCAGACCATCCACCTGGTTTATTTCTGGCGTTTCCAGATTACTATCCACCATCC 1670  
QY 1417 TGTGCATCTGAAATTCATATGTCTCTCTGGAAAGGCGAGCCCTGGTTGGCATTTATGGCAG 1476  
DB 1671 AATATATCTGAAGTTCAATATTTCTTTAGCCAAAGGACTCTCTGCTGGGAATTTATGGCAG 1730  
QY 1477 AAAAGGCTCTCTCTTCATACATACACAGTTTGACTTTTGTGGAGCTGTGGATGGCAGGAG 1536  
DB 1731 AAGAAACATTTCCACCTACACATCTCAGTTTGATTTTGTAAAACTAATGATGGCAACA 1790  
QY 1537 GCTCCTAACCCAGGAGGCGGAGCCTTAGAGGGGACCCCGCGCAGTCTCGGGGAACTGT 1596  
DB 1791 GCTGGTCAAGCAGGACTCCAAGGGCTCTGATGATACACAGCACTCCCTCGGAACCTGAT 1850  
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTTCATCCAGTATTTGGATTTCAGGAATCTGGCATT 1656  
DB 1851 CTTAACTTCGCTTCAGGAGACAGGTTTTCATAGAGTATATGGATCAAGGACCTTGGTATCT 1910  
QY 1657 GGCCTTTTACATGACGGAAGGAGTCAAGAGTGGTTTCTCTTCCACTGCCATTGA 1716  
DB 1911 GCGCTTTTACATGATGGAAGGAGATGGCAAGTATTCGTGTTAACTACAGCAATGA 1970  
QY 1717 GTCGGTGGATAACTGCCCCCAGCAACTGCTATGGCAATGGTACTGATCTCTGGGACCTG 1776  
DB 1971 AATAATGGATGACTGTTCAACCAATTGCAATGGAATGGAGAGTATCTCTGGCCATTG 2030  
QY 1777 CCACTGCTTCTGGGTTTCTCTGGGCCCCGACTGTGGCAGAGCCTCTGCCCCGCTCTG 1836  
DB 2031 TCATTGTTTCCAGGATTCCTTGGACCTGACTGTGCTAGAGATTCTGCTGCTGTGCTGTG 2090  
QY 1837 TAGCGGAAATGGCAATACATGAAAGGAGAGTCTTGTGCCACAGTGGCTGGAAGGCGC 1896  
DB 2091 TGGTGGGAATGGAGAATACGAGAAAGGACACTGTGTCTGCCGGCATGGCTGGAAGGGCC 2150  
QY 1897 TGAGTGCATGTGCCCAACCAACCACTGATGATGATGCTGCGCTGCAGCAACCATGGCACCTG 1956  
DB 2151 AGAGTGTGACGTTCCGGAAGAACAAATGCTATGATCCAACTGCTTGGCCACGGCACCTG 2210  
QY 1957 CATCAGGGCACTGCTCATCTGCAACCTGGCTACAGGGCGAGAGCTGTGAGGAAGTGA 2016  
DB 2211 CATCATGGGAGTCTGCTGCTGTGTCAGGATACAAAGGAGAAATATCGAGGAAGAGGA 2270  
QY 2017 CTGCATGGACCCCATGTTTCAGGCGCGGGTGTCTGCGTGAAGGCGAATGCCATTGCTT 2076  
DB 2271 CTGCCTAGACCCCAATGTGTTCCAACTGATGCTGCTGTAAAGGAGAAATGTCACTGTTT 2330  
QY 2077 TGTGGATGGGAGGACCAACTGCGAGACCCCGGCTTTGACGCTGTGACCCAGCTGGACTGG 2136  
DB 2331 TACTGGCTGGGAGGAGTTAACTGTGAAACACCACTTCTCTGTATGTCAAGAGCAGTGTCTC 2390  
QY 2137 AGCCACCGGAACCTTCTCCCGGACACCGGGCTTTGACGCTGTGACCCAGCTGGACTGG 2196  
DB 2391 AGGACACCGGAACCTTTCTTCTGGACGCTGGAGTATGCAGCTGTGATCCCAAGTGGACAGG 2450  
QY 2197 ACACGACTGTTCTATCGAGATCTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGG 2256  
DB 2451 ATCTGACTGCTCAACAGAGCTGTGTACCATGGAGTGTGGTAGCCATGGAGTCTGCTCAAG 2510  
QY 2257 GGGCACCTGCGCTGCGAGGATGGCTGGATGGGGGAGCCCTCGGACCAAGCGGGCTGCCA 2316

Db 2511 AGGAATTTGCCAGTGTGAAGAAGGCTGGGTAGGACCAACATGTAGGAACGCTCTGTCA 2570  
Qy 2317 CCCGGCTGTGCCGAGCATGGGACCTGCCGCGACGGAAGTGCAGTGCAGCCCTGGCTG 2376  
Db 2571 TTCTCANTGTACTGAGCATGGCCAAATGCAAAAGATGGAATAATGTAGTGTAGCCCTGGATG 2630  
Qy 2377 GAATGGCGAAACACTGCACCAATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCC 2436  
Db 2631 GGAGGCGACCACTGCACAAATGTCTCACTACTTAGAT---GCTGTCGAGATGGCTGCC 2687  
Qy 2437 TGGGTTGTCAATGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGGCTGTG 2496  
Db 2688 AGGGCTCTGCTTTGGAAATGGACGATGTACCTTGGATCAAAAATGGTTGGCACTGTGTGTG 2747  
Qy 2497 CCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCCCTGGGTGCACAG 2556  
Db 2748 TCAGGTGGGTTGGAGTGGGACAGGCTGCAATGTTGTCTAGGAAATGCTTTGTGGAGATAA 2807  
Qy 2557 CAAAGACAATGATGGAGATGGCTGGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCC 2616  
Db 2808 CTTGGACAATGATGGAGATGGTTAAACCGACTGTGTGGATCCTGACTGTTGTCAACAAAG 2867  
Qy 2617 CCTGTGCCATATCAACCCGCTGTGCTGGCTCCCTAACCCCTCTGGACATCATCCAGGA 2676  
Db 2868 CAACTGTTATATAAGTCCTCTCTGCCAGGGCTCACAGATCCTCTTGACCTCATTCAGCA 2927  
Qy 2677 GACACAGGTCCTGTGTCAACAGAGAACCTTACACTCCCTTCTATGACCGCATCAAGTTCC 2736  
Db 2928 AAGCCAAACTCTCTCTCTCAGCACACTTCAAGACTTTTTTATGATCGAATCAAATTCCT 2987  
Qy 2737 CGTGGGACGGGACAGCACCGCATATAATCCCCGGGGAGAACCCCTTTGATGGAGGGCATGC 2796  
Db 2988 CATTTGGCAAGGACAGTACTCATGTCTATTCCTCCTGAGGTGTCTATTTGACAGCAGCGCTGC 3047  
Qy 2797 TTGTGTTANTCGTGGCCAAAGTATGACATCAGATGGAAACCCCTGGTTGGTGTGAACAT 2856  
Db 3048 CTGTGTGATTCGAGGCCAAGTGTGGCCATAGATGGAATCCTCTAGTGGAGTGAATGT 3107  
Qy 2857 CAGTTTTGTCAATAAACCCCTCTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTGA 2916  
Db 3108 CAGTTTCTTGACCACAGTGATTATGGGTTTACCATCAGCCGGCAAGATGGAAGCTTGA 3167  
Qy 2917 CTTGGTGACAAAATGGCGGCATCTCCATCATCCTGCGGTTTCGAGCGGGCACCTTTTCATCAC 2976  
Db 3168 CCTCGTGGCCATCGGTGGCATCTCTGTCTCATCTTAATCTTCGACCGATCCCCCTTTCCCTGCC 3227  
Qy 2977 ACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTTCTTTGTCAATGGAAACCATCATCAT 3036  
Db 3228 TGAGAAGAGAAACACTCTGTTGCCCTTGGAAATCAGTTTATTGTGGTAGAGAAAGTCACCAT 3287  
Qy 3037 GAGACATGAGGAGAAATGAGATTCCCAGCTGTGACCTGAGCAATTTTGCCTCCGCCCCAACCC 3096  
Db 3288 GCAGAGAGTTGTATCAGACCCGCCCATCCTGCGATATCTCCAACTTTATCAGCCCCAAACCC 3347  
Qy 3097 AGTCGTCTCTCCATCCCCACTGACGTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCCAT 3156  
Db 3348 TATTGTGCTTCTTACCCGCTCACATCATTTGGAGGGTCTGTCCAGAGAGGGGAACTAT 3407  
Qy 3157 TGTGCCGGAAATTCAGGCTTTGCAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCT 3216  
Db 3408 TGTTCTGAGCTGCAGGTTGTACAGGAGGAAATTTCCCATTTCCCTCCAGCTTTGTGAGGCT 3467  
Qy 3217 GAGCTACCTGAGCAGCCGGACCCCTGGCTACAAATCTGTCTCTGAGGATCAGCCTCACCCA 3276  
Db 3468 GAGTTACCTGAGCAGCCGCACCCCTGGGTATAAAACCCCTGCTACGGATCCTTCTGACACA 3527  
Qy 3277 CCCGACCATCCCTTCAACCTCATGAAAGTGCACCTCATGTGTAAGCGGTGAGGGCCGCT 3336  
Db 3528 TTCAACGATTCCCGTAGGCATGATAAAAGTACACCTCACAGTAGCTGTGGAAGGGCGACT 3587  
Qy 3337 CTTCAGGAAGTGGTTCCGTGCAGCCCCGACCTGTCTTATTATTTGGAACAAGAC 3396

Db 3588 CACACAGAAGTGGTTTCCCGCCGCAATAAATCTTGTCTACACATTTGCTTGAACAAGAC 3647  
Qy 3397 AGACGTCTACAACACAGAAAGGTGTTGGGCTTTTCAGAAAGCCTTTGTTCCGTGGGTTATGA 3456  
Db 3648 CGATATCTATGGACAGAAAGGTTTGGGGCCTGGCAGAGGCTTTGGTATCTGTGGGATGA 3707  
Qy 3457 ATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAAACAACAGTGTGCAGGGCTA 3516  
Db 3708 ATATGAAACGTGCCCTGACTTTATCTCTGGGAGCAAGGACAGTCGTTTACAAAGGTTT 3767  
Qy 3517 TGAATTTGACGCGTCCAAGCTTGGAGGATGGAGCCTTAGACAAAACATCATGCCCTCAACAT 3576  
Db 3768 TGAGATGGATGCTTTCTAACCTAGGAGACTGGTCTTTTGAATAAGCATCACATTTTGAATCC 3827  
Qy 3577 TCAAAGTGGTATCCTGCAACAAGGGAATGGGAGAACCCAGTTTGTCTCTCAGCAGCCTCC 3636  
Db 3828 TCAAAGTGAATCATACATAAAGGGAATGGAGAAAATATGTTTCAATTTCCAGCAGCCCC 3887  
Qy 3637 TGTCAATTTGGAGCATCATGGGCAATGGGCGCCGGAGAGCATCTCTGCCCCCAGCTGCAA 3696  
Db 3888 AGTCATATCAACCATAAATGGGTAAATGGACACCAAGGAGTGTAGCCTGCACCAACTGCAA 3947  
Qy 3697 CGGCTTGTCTGACGGCAACAAGCTCCTGGCCCCCAGTGGCCCCCTCACCTGTGGCTCTGACGG 3756  
Db 3948 TGGCCCCAGCCACAACAACAACCTCTTTGCTCCTGTGCGCTTAGCTTCTGGCCCTGATGG 4007  
Qy 3757 GAGCCTCTATGTGGGTGATTTTCAACTACATTAGAAGGATCTTCCCTCTGGAATGTCTAC 3816  
Db 4008 CAGTGTGTATGTTGGCGACTTCAATTTTGTAAAGGAGAAATATTTCCCTCGGGAACCTCCGT 4067  
Qy 3817 CAACATCCTTAGAGCTGAGGAATAAAGATTTTCAACATAGTCAAGTCCAGCACACAATA 3876  
Db 4068 TAGTATTTTGAATAAAGCACA-----AGTCCTGTCTACAAATA 4106  
Qy 3877 CTACCTGGCCACAGACCCCATGAGTGGGGCCGTCTTCTCTTTCAGACAGCAACAGCCGGCG 3936  
Db 4107 CTATCTGGCTATGGACCCCTGTGTCTGAATCACTCTATCTATCAGACACCAATACTCGCAA 4166  
Qy 3937 GGTCCTTTAAATCAAGTCCACTGTGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGT 3996  
Db 4167 AGTCTACAAGTTGAAATCTCTTGTGGAGACGAAAGATCTGTCCAGAAATTTTGAAGTGGT 4226  
Qy 3997 TGCGGGGACAGGTGACCAAGTCCACTGTGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGG 4056  
Db 4227 GGCAGGAACCTGATGATCAGTGCCTTCCCTTTGACCAAGTCAATGTGGAGATGGTGGAG 4286  
Qy 4057 GGCCACAGAAAGCCACACTCACCAATCCCAGGGGTATTACAGTGGACAAGTTTGGGCTGAT 4116  
Db 4287 AGCATCGGAAGCTTCACTGAATAGCCCTCGAGGCATCACAGTTGATAGGCATGGATTTAT 4346  
Qy 4117 CTACTTCGTGATGGCACCATGATCAGACGCATCGATCAGAATGGGATCATCTCCACCT 4176  
Db 4347 TTACTTTGTGGATGGACTATGATTCGCAAAATTTGATGAGAATGCTGTGATCACAACCTGT 4406  
Qy 4177 GCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGTGATTTCTGTCTATGGATAT 4236  
Db 4407 AATCGGCTCAAATGGTCTGACTTCCACACAACCACTGAGCTGTGACTCAGGAATGGACAT 4466  
Qy 4237 TTCCCAGGTAAAGACTGGAGTGGCCCAAGACTTAGCCATCAACCCAAATGGACAACCTCACT 4296  
Db 4467 CACTCAGGTGCGATTAGAGTGGCCAAACAGACCTTGCAGTAAATCTTATGGACAATTCATT 4526  
Qy 4297 TTATGTCTCGACAACAATGTGGTCCCTGCAAAATCTCTGAAAAACCAACAGGTGCGCATTTGT 4356  
Db 4527 GTATGTCTTGGATAACAACATTTGTGTGCAAAATTTCTGAGAACAGGCGTGTTCGGATCAT 4586  
Qy 4357 CGCCGGGAGGCCCATGCACTGCCAGGTCCCTGGGCATTTGACCATTGACCATTCTCTGCTAAGCAAGGT 4416  
Db 4587 CGCAGGACGCCCCATTCACCTGCCAGGTGCCAGGCATCGATCATTTTCTCTGTCAGCAAGGT 4646  
Qy 4417 GGCCATCCACGCCAACCTTGAGTGCAGCCACCGCTTTGGTGTTTTACAAATGGGGTCTCT 4476  
Db 4647 AGCAATTCACCTCCACTCTAGAGTCAGCGAGGGCCATCAGTGTCTCCACAGCGGGCTGCT 4706

QY 4477 GTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCCACACTAGTGG 4536  
 Db 4707 CTTATAGCTGAACAGACGAGAGGAAAGTAAACCGCATTCAGCAAGTAAACCAATGG 4766  
 QY 4537 AGAGATCTCACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAATAATGATGCCAACTG 4596  
 Db 4767 GGAGATCTACATCATCGCTGGTCCCCCACTGACTGTGACTGCAAAATGATCCAAACTG 4826  
 QY 4597 TGATTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAGTTAAATACCCCATCTTC 4656  
 Db 4827 TGAATGTTTTTCAGGTGATGGTGGCTATGCCAAGATGCAAGATGAAAGCCCCCTTCCTC 4886  
 QY 4657 CTTGGCTGTGTGTGCTGATGGGAGCTCTACGTGCCCGACCTTGGGAACATCCGAATTCG 4716  
 Db 4887 CTTAGCAGTGTGCGCTGTATGGAACCTCTATGTGGCAGACCTCGGAATGTTGGAATTCG 4946  
 QY 4717 GTTTATCCGGAAGAACAGCCCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTCACC 4776  
 Db 4947 TACCATCAGCAGGAACCAAGCCCACTGAAATGACATGAACATTTATGAGATTGCTTCACC 5006  
 QY 4777 AATTGACCAGGAGCTCTATCTGTTGTATACCAACCGGCAAGCACCTGTACACCCAAAGCCT 4836  
 Db 5007 CGCTGATCAGGAACCTGTACAGTTCACTGTAAATGGAACCCACCTACACACCCCTGAACTT 5066  
 QY 4837 GCCCACAGGAGACTACCTGTACAACTTCACTACACTGGGACCGGACATCACACTCAT 4896  
 Db 5067 GATAACAAGGACTATGTTTATAACTTCACTTACCTACAATTTCTGAAGGTGACTTGGGCGGAT 5126  
 QY 4897 CACAGACAAACAATGGCAACATGGTAAATGTCCGCGGAGACTCTACTGGGATGCCCTCTG 4956  
 Db 5127 TACCAGCAGCAATGGCAATTCAGTGACATTCGCGCTGATGCAGGCGGAATGCCGCTATG 5186  
 QY 4957 GCTGGTGGTCCAGATGGCCAGGTGTACTGGTGACCATGACATACCATGGCAATTCGCGCCT 5076  
 Db 5247 AAGAGTGTAGCCCAAGGCTATAATCCGCGCTTAATGACCTATCCAGGAAACACAGGGCT 5306  
 QY 5077 TCTGGCAACCAAGCAATGAAACCGATGGACCAACATTTTATGATGACGACAGCTTTGG 5136  
 Db 5307 TCTGGCTACCAAAAGTAACGAAATGGATGGACCAACCGTTTATGATGATGACCCCGAGGG 5366  
 QY 5137 CCGCCTGACAAATGTGACCTTCCCTACTGCGCAGGTGAGCAGTTTCCGAAGTGATACAGA 5196  
 Db 5367 ACACCTGACCAATGCAACGTTTCCCACTGGAGAGGTGAGCAGCTTCCACAGTGACCTGGA 5426  
 QY 5197 CAGTTCAGTGATGCCAGGTAGAGACCTCCAGCAAGGATGATGTACCCATAACCAACCAA 5256  
 Db 5427 GAAGCTGACAAAGTGGAGCTAGATACITCCCAACCGTGAAATGTCTCATGTCAACCAA 5486  
 QY 5257 CCTGTCTGCCCTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTA 5316  
 Db 5487 CTTGACGGCAACTAGTACCATATATATTTTAAACAAGAAATACTCAAAAGTACCTATCG 5546  
 QY 5317 CATCGGGCCGATGGCTCCTTGGCGCTGCTGGCCCAACGGCATGGAGGTGGCGCTGCA 5376  
 Db 5547 GGTGAATCCAGATGGTTCCTGCGTGTACTTTTGGCAGCGGGATGGAGATCGGCTCAG 5606  
 QY 5377 GACTGAGCCCCACTTGTGGTGGCAACCGTCAACCCCAACCGTGGGCAAGAGGAATGTAC 5436  
 Db 5607 CTCAGAGCCCCACATCTCGCAGGGGAGTCAACCCCTACCCCTGGGCAAAATGCAACATCTC 5666  
 QY 5437 GCTGCCCCATCGAACAGGCCCTCAACCTGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCG 5496  
 Db 5667 ATTGCCCCGAGAGCACAATGCAAACTCATCGAGTGGCGGAGAGGAGGCAAAACAA 5726  
 QY 5497 GGGCCAGGTCACTGTCTTTGGCGCCGGTGGGCTGCACAAACCGAAATCTCCTATCTCT 5556  
 Db 5727 AGGCAATGTTTCGGCTTTTGAAGGAGGCTGAGGGCCCAACACAGAAACCTACTCTCCAT 5786

QY 5557 GGACTTTGATCGCGTAAACACGCACAGAGAAAGATCTATGATGACCAACCGAAGTTCCACCCT 5616  
 Db 5787 AGATTTTGATCATATAACCCGCACAGGAAAGATCTATGATGACCATCGAAATTCACCCT 5846  
 QY 5617 TCGGATTTCTGTACGACAGCGCGGGCGGCCAGCTCTGGTCAACCCAGCAGCAGGCTGAA 5676  
 Db 5847 TCGAATTTCTTATGACAGACTGGCGCACCCCACTTCTGGTCTCCTGTAAAGCAGATATAA 5906  
 QY 5677 TGGTGTCAACGTGACATFACTCCCTGGGGGTTCATATGCTGGCATCCAGAGGGGCATCAT 5736  
 Db 5907 TGAAGTGAACATCACATATTCACCTTCGGGATTTGGTGACGTTTATTCAAGAGAGAACGTG 5966  
 QY 5737 GTCTGAAAGAAATGGAATPACGACAGCGCGGCGCATCACATCCAGGATCTTCGCTGATGG 5796  
 Db 5967 GAATGAAAAAATGGAATATGACAGAGTGGGAAAAATTTTCAAGAACTTGGGCTGATGG 6026  
 QY 5797 GAAGACATGGAGCTACACATACTTTAGAGAAAGTCCCATGGTGTCTGTACTACACAGCCAGAG 5856  
 Db 6027 GAAAAATTTGGAGCTATACCTACTTTAGAAAAATCTGTGATGCTTCTCTACACAGCCAGCG 6086  
 QY 5857 GCAGTATATCTTTGAGTTCGACAAGAATGACCGCCTCTCTTCTGTGACGATGCCCAAGCT 5916  
 Db 6087 CGGTTACATCTTTGAGTATGACCAATCAGATGCGCTGTGTGAGTTACCATGCCTAGCAT 6146  
 QY 5917 GCGCGGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCC 5976  
 Db 6147 GGTGCGCCACAGCTTACAAACCATGTCTTTCAGTGGGCTACTACCGTAATATCTACACCCC 6206  
 QY 5977 CCCTGAGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGACCTCTTTCACAC 6036  
 Db 6207 ACCGACAGTAGCACTTCTTTTATCCAAGACTATAGTCGAGATGGCCGATGCTACAGAC 6266  
 QY 6037 CTTCTACCTGGGACTGGCCGACGGGTGATATACAAGTATGGAAACTGTCAAAGCTGGC 6096  
 Db 6267 CCTGCATCTGGGACAGGCGCAGAGTCTTATACAAGTACACCAAGCAAGAGGCTTTC 6326  
 QY 6097 AGAGACGCTCTATGACACCAACAGGTGAGTTTCACTATGACGACGCGCAGGATGCT 6156  
 Db 6327 TGAGGTTCTCTATGATACCACTCAGGTACATTAACATATGAAGACTTCTTCTGGAGTAT 6386  
 QY 6157 GAAGACCATCAACCTACAGAATGAGGCTTCACTGACCATCCGCTACCGTCAGATTGG 6216  
 Db 6387 TAAAGACAATACACCTGATGATGACGAGTTCATCTGCACAAATCAGATACAGGCAACAGG 6446  
 QY 6217 GCCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCAATGGTCAACGCCGTTT 6276  
 Db 6447 ACCTCTTATGGACGCCAGATTTTTCAGATTCAGTGAAGAGGCTTGTGAATGCACGTT 6506  
 QY 6277 TGACTACAACTATGACAAACAGCTTCCGCTGACAGCATGAGGCTGTGTGATCAACGAGAC 6336  
 Db 6507 CGACTACAGCT--ACAACAATTTCCGAGTCAAGAGCATGCAAGCTGTAAATCAATGAAAC 6563  
 QY 6337 CCCACTGCCAATTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGG 6396  
 Db 6564 CCTTTGCTATAGATCTTTACCGATATGTGTGATGTCTCTGGCAGAACAGAGCAGTTTGG 6623  
 QY 6397 GAAGTTTGGTGTCACTTACTATGACATTAACCCAGATCATCACCAAGCTGTGATGACCCA 6456  
 Db 6624 AAAATTGAGTGAATTAATTACGATTTAAATCAGGTCTATACTACTACAGTGTGATGAAACA 6683  
 QY 6457 CACCAAGCATTTTGTGATGATGATGGCAGGATGAAGGAGTGCAGTATGAGATCTTCCGCTC 6516  
 Db 6684 CACCAAAATCTTCAGTGCCAATGGACAAAGTCAATGAAGTCCCAATGAATCCTAAAGGC 6743  
 QY 6517 GCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGCGGAGTAGTGAAGAAG-GAGC 6575  
 Db 6744 AATTGCCTACTGGATGACCATTTCAATATGATAATGTGGCCGACATGGTAAATGTGACA 6803  
 QY 6576 TGAAGGTAGGACCTACGCCAATACCACTCGCTACTCTCTATGATGATGATGATGATGATG 6635  
 Db 6804 TAAGGTAGGAGTAGATGCCAATATAACAAGGTACTTCTATGAATACGATGCTGATGGGC 6863  
 QY 6636 AGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATG 6695

Db 6864 AACTCAGACTGTTCTGTAAATGACAAAACCCAGTGGCGTTATAGTTACGATCTGAATG 6923  
Qy 6696 GGAACCTGCACCTTACTAGCCCTGGGAACAGTGCACGGCTCACACCACHACGGTATGACA 6755  
Db 6924 GAAACATCAACCTCTTAAGCCATGGGAAGAGTGCCTTACTCCTCTCCGATATGACC 6983  
Qy 6756 TCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGCTTCC 6815  
Db 6984 TCCGAGACCGCATCACAGATTAGGAGAAATTCAGTATAAAATGGATGAAGATGGCTTTC 7043  
Qy 6816 TGAGGCAGCGGGCGGTGATATCTTTTGTAGTACAACCTCAGCTGGCCTGCTCATCAAGGCCT 6875  
Db 7044 TGAGGCAGAGGGGAAATGATATTTTGAATATAATTTCTAATGGCCTGCTGCAGAAAGCCT 7103  
Qy 6876 ACAACCGGGCTGGCAGCTGGAGTGTCAAGTACCACTCAGCTGGCCTGCTCATCAAGGCCT 6935  
Db 7104 ACAATAAGGCTTCTGGCTGGACTGTGCAGTATTACTATGATGGCTTGGCGACGTGTCG 7163  
Qy 6936 CCAGCAAGAGCAGCCACAGCACACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACT 7055  
Db 7164 CGAGTAAGTCCAGCCTAGGGCAGCACCTTCAGTTCTTTTATGAGACCTTACCAACCCCA 7223  
Qy 6996 CCAAGGTCACCCACCTGTACAAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACT 7055  
Db 7224 TAAGAGTTACTCATTTGTACAACACACACAAGCTCGGAGATTACATCTCTGTATTATGATC 7283  
Qy 7056 TGCAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGATGAGTTTACATAGCTTGTG 7115  
Db 7284 TCCAGGTCACCTTATTGCCATGGAGTTAAGCAGTGGTGAAGAATATTATGTAGCCTGTG 7343  
Qy 7116 ACAACATCGGGACCCCTCTTGTCTCTTTTAGTGAACAGGTTTGATGATCAAGCAAAATCC 7175  
Db 7344 ATAATACAGGTACCCCACTAGCTGTGTTTCAGCAGCCGAGGTGAGTCAATAAGGAGATAC 7403  
Qy 7176 TGTACACAGCCTATGGGGAGATCTACATGGATACCAACCCCAACTTTTCAGATCATATAG 7235  
Db 7404 TATACACACCTTATGGCGATATCTATCATGACACTTACCCTGACTTTCAGGTCAATTATG 7463  
Qy 7236 GCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGCGCGGAGATT 7295  
Db 7464 GTTTTCATGAGGACTCTATGATTTCTTACTAAATTTAGTGACCTGGGGCAAGGGATT 7523  
Qy 7296 ATGATGTGCTGGCCGACCGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACCTTAGTA 7355  
Db 7524 ATGATGTGTGCTGGCAGATGGACAACGCCTAATCATCATATGGAACAGTT--GA 7580  
Qy 7356 GCAGCAACGTCAATGCCCTTTTAATCTCTATATGTTCAAAAACAACCAACCCCATCAGCAACT 7415  
Db 7581 ACCTCCTTCTTAAACCAATTCACCTCTACTCCTTTTGAAAATAAATACTACCCAGTTGGCAAAA 7640  
Qy 7416 CCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTTGGATTCC 7475  
Db 7641 TTCAAGATGTGCAAAAGTATACCCACAGACATCAGAAGTTGGTTGGAGCTATTTGGTTTCC 7700  
Qy 7476 AGCTACACAACGTGATCCCTGGTTATCCCAACCCAGACATGGATGCCATGGAACCCCTCCT 7535  
Db 7701 AATTACACAATGTACTACCTGGATTTCCTCAACCTGAATTAGAAAAATTAGAAATTAACCTT 7760  
Qy 7536 ACGAGCTCATCCACACACAGATGAAAAACGAGGAGTGGGACAAACAGCAAGTCTATCCTCG 7595  
Db 7761 ACGAGCTTCTACGGCTTCAGACAAAAAATCAAGATGGGATCCTGGAAAGACTATCCTGG 7820  
Qy 7596 GGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCTTTGTACCTTAGAACGGTTTGTACC 7655  
Db 7821 GCATTCAGTGTGAACCTCCAGAAACAGCTCAGGAATTTTCATTTCTTGGACCAACTACCTA 7880  
Qy 7656 AGCTCTATGGTCCCAAAATCACACAGCTGCCAGCAGGCTCCAAAGACCAAGAGTTTGTGAT 7715  
Db 7881 TGACTCCCCGATACAATGATGGACGGTGCCTTGAAGGAGGGAAGCAACCAAGGTTTGTG 7940  
Qy 7716 CCAGCGGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCA 7775

Db 7941 CTGTCCCTTCTGTTTTTTGGGAAAGGTATAAAATTTGCCATCAAGGATGGCATAGTAACAG 8000  
Qy 7776 CAGACATCATAGTGTGGCCAATGAGGATGGGCGAAGGTTGCTGCCATCTTGAACCATG 7835  
Db 8001 CTGATATTATAGGAGTAGCCAATGAAGATAGCAGGCGGCTTGCTGCCATTCTCAATAATG 8060  
Qy 7836 CCCACTACCTAGAGAACCTGCACTTCAACATTGATGGGTGGATACCCATTACTTTGTGA 7895  
Db 8061 CCCATTACCTGGAAAACCTACATTTTACCATAGAGGGAGGACACTCACTACTTCATTA 8120  
Qy 7896 AACCAAGACCTTCAGAAGGTGACCTGGCCATCTCTGGCCTCAGTGGGGCGGCGGAACCC 7955  
Db 8121 AGCTTGGGTCTCTGGAGGAAGACCTGGTGTCTATCGGTAACACTGGGGGAGCGGATTC 8180  
Qy 7956 TGGAGAATGGGGTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTA 8015  
Db 8181 TGGAGAATGGTGTCAATGTCACTGTGTCCAGATGACTTCTGTGTGAATGGGAGGACTA 8240  
Qy 8016 GACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAACACACACGCTACGGGA 8075  
Db 8241 GACGGTTGCAGATATTTCAGCTCCAGCATGGAGCCCTGTGTCTCAACATCCGGTATGGGA 8300  
Qy 8076 CAACGTTGGATGAGGAGAAGGCACGGGTCTCTGGAGCTGGCCCGGAGAGAGCCGTGCGCC 8135  
Db 8301 CAACGTGCGAAGAGGAAAAGAAATCACGTGTTGGAGATTGCCAGACACAGCGCAGTGGCCC 8360  
Qy 8136 AAGCGTGGCCCCGCGAGCAGCAGACACTGCGGGAAGGGAGGAAAGGCTGCGGGCCTGGA 8195  
Db 8361 AGCCCTGGACTAAGGAACAAAGAGGCTGCAAGAGGGGGAAGAGGGGATTAGGGCATGGA 8420  
Qy 8196 CAGAGGGGAGAAAGCAGCAGGTTGCTGAGCACAGGGCGGGTGCAGAGGCTACGACGGCTTTT 8255  
Db 8421 CAGAAGGGGAAAAGCAGCAGCTTTTGAGCAGCTGGCGGGGTACAAGTTACGATGGGTATT 8480  
Qy 8256 TCGTGATCTCTGTCGAGCAGTACCAGAACTGTTCAGACAGCGCCCAACACATCCACTTCA 8315  
Db 8481 TTGTTTTGTCTGTTGAGCAGTATTTAGAACTTTCTGCACAGTGCCTAATAATTACCTTTA 8540  
Qy 8316 TGAGACAGAGCGAGATGGGCCGAGGTGACAGAGA 8350  
Db 8541 TGAGACAGAGCGAAAATAGGCAGGAGGTAAACAAAA 8575

RESULT 13

ABK92230

ID ABK92230 standard; DNA; 12879 BP.

XX ABK92230;

XX AC ABK92230;

DT 15-AUG-2002 (first entry)

XX Prostate cancer-associated DNA sequence #116.

DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;

XX gene therapy; gene; ds.

XX Mammalia.

OS WO200230268-A2.

PN 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.

XX 13-OCT-2000; 2000US-00687576.

PR 08-DEC-2000; 2000US-00733288.

PR 08-DEC-2000; 2000US-00733742.

PR 24-JAN-2001; 2001US-0263957P.

PR 16-MAR-2001; 2001US-0276791P.

PR 16-MAR-2001; 2001US-0276888P.

PR 06-APR-2001; 2001US-0281922P.

PR 24-APR-2001; 2001US-0286214P.

PR 30-APR-2001; 2001US-00847046.

PR 04-MAY-2001; 2001US-0288589P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX WPI; 2002-471335/50.  
DR P-PSDB; ABG61913.  
XX  
PT Detecting a prostate cancer-associated transcript in a cell in a patient,  
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
PT by determining if prostate cancer-associated genes are expressed in a  
PT prostate tissue.  
XX  
PS Claim 22; Page 394-397; 436pp; English.  
XX  
CC The present invention relates to methods of detecting a prostate cancer-  
CC associated transcript in a cell from a patient. The method comprises  
CC contacting a biological sample from the patient with prostate cancer-  
CC associated polynucleotides (designated PC genes) that selectively  
CC hybridize to a sequence that is at least 80% identical to them. The  
CC prostate cancer-associated polynucleotide sequences are differentially  
CC expressed in prostate tumour tissue or in prostate cancer and are derived  
CC from the tissues of various organisms such as humans or other mammals  
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
CC associated genes are useful for diagnosing or treating prostate cancer,  
CC as well as for identifying modulators of prostate cancer or agents that  
CC inhibit prostate cancer. The nucleic acid sequences are particularly  
CC useful in gene therapy, as a vaccine or in antisense applications.  
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide  
CC sequences  
XX  
SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Query Match 39.0%; Score 3258; DB 6; Length 12879;  
Best Local Similarity 65.5%; Pred. No. 0;  
Matches 4893; Conservative 0; Mismatches 2515; Indels 66; Gaps 6;  
QY 880 CGGGCACTTCTCTTCAAGCCTGGA--GGACCTTCCCGCTCTTCTGCACCACATCACC 936  
DB 838 CAGGCATTCCCTGTTCAACATGGATCTGGTTCTCTGCGATCTTCAAGTGCAGCCAGTCA 897  
QY 937 AGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCGCCCGACCCCTGCCCGCAG 996  
DB 898 GAACCTACCTCTGACATCCATACCGTGTACTGCCCCCTCCAGGCTCTTCTCCTCGAAG 957  
QY 997 CACCTTGGCCCGCGGCTTTAACTCAAGAGCCCTTCAAGTACTGTAACTGGAAGTG 1056  
DB 958 CACCTTTTCCCGACCTGCTTTTACCTTTTAAACAACTTACAGGTGCTGCAACTGGAAGTG 1017  
QY 1057 GCAGCCCTGAGGCCATCGTCATCTCAGCCACTCTGGTCACTCTGCTGGCATACTTTGT 1116  
DB 1018 CACAGCATTTGAGCGCCACTGCAATCAGTGTACTTGGCCCTTGTACTAGCCTATGTGAT 1077  
QY 1117 GGCCATGCACCTGTTGGCCCTAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTATGA 1176  
DB 1078 TGCAGTGCATTTGTTCGGCCTGACTTGGCAGTTGCAACCACTTACAGGTGAGAGAGCTGTATGC 1137  
QY 1177 GATCAGGAGGACAGCCAGCAGTGTGGCTGTGCGCAACCGACGTCCTCCCTATACCCCTC 1236  
DB 1138 AATGGAGTTAGCAAAGGGAACAGGGGACCGAGTCCATGGACACTACTTACTCTCCAAT 1197  
QY 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGGAACCAAGAAAGGCC 1296  
DB 1198 TGGAGGAAAAGTTTCTGATAAATCAGAAAAAAG----- 1232  
QY 1297 CAGTAGTTTCTTCCAGAGGACAGTTTTCATAGATTTCTGGAGAAATGTATGTGGAGAGCG 1356  
DB 1233 -----TGTTCAGAAAGGACGGCGGATAGACACTGGAGAGTTGACATTGGTGACA 1284  
QY 1357 AGCCTCCAGAGATTCCTCTCGCACTTCTTCTGGAGATCTCAAGTCTTCATAGACCATCC 1416

DB 1285 GGTCAATGCAGACCAATTCACACTGGTTTATTCTGGCGTTTCCAGATTACTATCCACCATCC 1344  
QY 1417 TGTGCATCTGAAATTCATGTGTCTCTGGAAAGCAGCCCTGGTTGGCATTTATGGCAG 1476  
DB 1345 AATATATCTGAAGTTCATATATTCTTTAGCCAAAGACTCTCTGCTGGGAATTTATGGCAG 1404  
QY 1477 AAAAGGCCTCCCTTCCATACATACACAGTTTGTGAGCTGTGGAGCTGTGGATGGCAGGAG 1536  
DB 1405 AAGAAACATTCACCTACACATCTCAGTTTGTAAACTAATGATGGCAACA 1464  
QY 1537 GCTCCTAACCCAGGAGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGT 1596  
DB 1465 GCTGGTCAAGCAGGACTCCAAGGCTCTGATGATACACAGCACTCCCTCGGAACCTGAT 1524  
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTTCATCCAGTATTTGGATTTCAGGAATCTGGCATT 1656  
DB 1525 CTAACTTCGCTTCAGGAGACAGGTTTCATAGAGTATATGGATCAAGGACCTTGGTATCT 1584  
QY 1657 GCCTTTTACAATGACGAAAGGAGTCAAGTGGTTTCTCTTCTCACCCTGCCATTGA 1716  
DB 1585 GCGGTTTACAATGATGGAAGAAAGATGGAGCAAGTATTCGTGTTAACTACAGCAATTGA 1644  
QY 1717 GTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTACTGCTATCTCTGGGACCTG 1776  
DB 1645 AATAATGGATGACTGTTCAACCAATTGCAATGCAATGGAGATGATATCTCTGGCCATTG 1704  
QY 1777 CCACTGCTTCTGGGTTTCTGGGCCCCGACTGTGGCAGAGCCCTCTGCCCCCGTCTCTG 1836  
DB 1705 TCATTGTTTCCAGGATTCCTTGGACCTGACTGTGCTAGAGATTCTCTGCCCTGTGCTGTG 1764  
QY 1837 TAGCGGAAATGGCAATACATGAAAGGCAAGATGCTTGTGCCACAGTGGCTGGAAGGCGC 1896  
DB 1765 TGGTGGGAATGGAAATACGAGAAAGGACACTGTGTCTGCCGCAATGGCTGGAAGGGCC 1824  
QY 1897 TGAGTGGATGTGCCACCAACCAAGTGTATGATGTGGCTGCAGCAACCATGGCACCTG 1956  
DB 1825 AGAGTGTGACGTTCCGGAAGAAACAATGCAATGATCCAACTGCTTTGGCCACGGCACCTG 1884  
QY 1957 CATCACGGGCACCTGCACTCTGCAACCTTGGCTACAGGGCGAGAGCTGTGAGGAAGTGA 2016  
DB 1885 CATCATGGGAGTCTGCACTCTGTGTGCCAGATACAAAGGAGAAATATCGGAGGAAGAGGA 1944  
QY 2017 CTGCATGGACCCCACTGTTTCAAGCCGCGGTGTCTGCGTGAGAGGCAATGCCATTGCTT 2076  
DB 1945 CTGCCTAGACCCCAATGTGTTTCAACCACTGSCATCTGTGTAAAGGAGAAATGCTACTGTT 2004  
QY 2077 TGTGGGATGGGAGGCACCAACTGCGAGACCCCGGCGCCACATGTTAGACCAGTGTTC 2136  
DB 2005 TACTGGCTGGGAGGAGTTAACTGTGAAACACCACTTCTCTGTATGTAAGAGCAGTGTCT 2064  
QY 2137 AGCCACGGAACCTTCTCCCGGACACCGGGCTTTTCAGCTGTGACCCCAAGCTGGACTGG 2196  
DB 2065 AGGACACGGAACCTTCTCTGCGACGCTGGAGTATGCACTGTGATCCCAAGTGGACAGG 2124  
QY 2197 ACAGCACTGTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGG 2256  
DB 2125 ATCTGACTGCTCAACAGAGCTGTGTACCATGGAGTGTGGTAGCCATGGAGTCTGCTCAAG 2184  
QY 2257 GGGCACCTGCCCTGCGAGGATGGCTGGATGGGGGCGAGCTGGGACCGAGCGGCTGCCA 2316  
DB 2185 AGGAATTTGCCAGTGTGAAGAGGCTGGGTAGGACCAACATGTGAGGAACGCTCCTGTCA 2244  
QY 2317 CCCGCGCTGTGCCGAGCATGGGACCTGCGCGCGACGGCAAGTGGCAGTGCAGCCCTGGCTG 2376  
DB 2245 TTCTCATTTGACTGAGCATGGCCCAATGCAAAAGATGGAATGTAAGTGTAGCCCTGGATG 2304  
QY 2377 GAATGGCGAACACTGCACCATCGCTCACTATCTGGATAGGTTAGTTAAAGAGGGTTGCC 2436  
DB 2305 GGAGGGCGACACTGCACAATTTGCTCACTACTTAGAT---GCTGTCCGAGATGGCTGCC 2361  
QY 2437 TGGGTTGTGAATGGCAACCGGAGATGTACTCTTAGACCTGAATGGTTGGCACTGCGTCTG 2496  
DB 2362 AGGGCTCTGCTTTGGAATGGACGATGTACCTTGGATCAAAATGGTTGGCACTGTGTGTG 2421





QY 6877 CAACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCCTGGGCGCGCGTGTG 6936  
Db |||||  
QY 6937 CAGCAAGAGCAGCCACAGCCACCACCTGCGAGTTCTTCTATGCAGACCTGACCAACCCAC 6996  
Db |||||  
QY 6997 CAAGGTCACCCACCTGTACAACCACCTCCAGCTCTGAGATCACCTCCCTCTACTACGACTT 7056  
Db |||||  
QY 6957 AAGAGTTACTCATTTGTACAACCACACAAGCTCGGAGATTACATCTCTGTATTATGATCT 6957  
QY 7057 GCAAGGACACCTCTTTGCCATGGAGCTGAGCTGAGCAGTGGTGTGATGAGTTTACATAGCTTGTGA 7116  
Db |||||  
QY 6958 CCAAGGTCACCTTATTGGCATGGAGTTAAGCAGTGGTGAAGATATTATGTAGCCTGTGA 7017  
QY 7117 CAACATCGGGACCCCTCTTGCTGTCTTTAGTGGBAACAGGTTTGTATGATCAAGCAATCCT 7176  
Db |||||  
QY 7018 TAATACAGGTACCCCACTAGCTGTGTTTCAGCAGCCGAGGTCAGGTCATAAAGGAGATACT 7077  
QY 7177 GTACACAGCCTATGGGGAGATCTACATGGATACCAACCCCAACTTTCAGATCATCATAGG 7236  
Db |||||  
QY 7237 CTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCACATGGGCCGCGAGATTA 7296  
Db |||||  
QY 7138 TTTTCATGGAGACTCTATGATTTCTTACTTAATTAGTGACCTGGGGCAAAGGGATTA 7197  
QY 7297 TGATGTGCTGGCCGACGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACTTAGTAG 7356  
Db |||||  
QY 7357 CAGCAACGTCATGCCCTTTTAACTCTCTATATGTTCAAAAACCAACCCCACTCAGCAACTC 7416  
Db |||||  
QY 7417 CCAGGACATCAAGTGTCTTCATGACAGATGTTTAACAGTGGCTGTCTCACCTTTGGATTCCA 7476  
Db |||||  
QY 7315 TCAAGATGTTGCAAGTATACCAACAGACATCAGAAAGTGGTTGGAGCTATTGGTTTCCA 7374  
QY 7477 GCTACACAACGTTGATCCCTGGTTATCCCAACCCAGACATGGATGCCATGGAAACCCCTCCTA 7536  
Db |||||  
QY 7375 ATTACACAATGTACTACCTGGATTTCCTCAACCTGAATTAGAAATTTAGAAATTAACCTTA 7434  
QY 7537 CGAGCTCATCCACACACAGATGAAAACGCAGGAGTGGGACCAACAGCAAGTCTATCTCTCGG 7596  
Db |||||  
QY 7435 CGAGCTTCTACGGCTTCAGACAAAACCTCAAGAGTGGGATCCTGGAAAGACTATCCTGGG 7494  
QY 7597 GGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTTGACCA 7656  
Db |||||  
QY 7495 CATTCAGTGTGAACCTCCAGAAACAGCTCAGGAATTTCAATTCCTGGACCAACTACCTAT 7554  
QY 7657 GCTCTATGGCTCCACAATCACAGCTGCCAGCGGCTCCAAAGACCAAGAGTTTGCATC 7716  
Db |||||  
QY 7717 CAGCGGCTCAGTCTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGAACCA 7776  
Db |||||  
QY 7674 TGTCCCTTCTGTTTTTGGGAAAGGTATATAAAATTTGCCATCAAGGATGGCATAGTAACAGC 7674  
QY 7777 AGACATCATCAGTGTGGCCATGAGGATGGCGAAGGTTGTGCCATCTTGAACCATGC 7836  
Db |||||  
QY 7675 TGATATTATAGGAGTAGCCAATGAAGATAGCAGCGGCTGTGCTGCCATTCCTCAATAATGC 7734  
QY 7837 CCACCTACCTAGAGAACCTGCACCTTACCATTGATGGGTGGATACCCATTACTTTGTGAA 7896  
Db |||||  
QY 7735 CCATTACCTGGAAAACCTACATTTTACCATAGAGGGGAGGAGACACTCACTACTTCATTAA 7794  
QY 7897 ACCAGACCTTCAGAAAGTGACCTGGCCATCCTGGGCCCTCAGTGGGGCGGCAACCCCT 7956  
Db |||||  
QY 7795 GCTTGGGTCTCTGGAGGAAGACCTGGTGTCTCATCGGTAACACTGGGGGAGGCGGATTCCT 7854

QY 7957 GGAGAAATGGGGTCAACGTCACCTGTGTCCCAGATCAACACAGTACTTAATGGCAGGACTAG 8016  
Db |||||  
QY 8017 ACCTTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAACACACAGCTACGGGAC 8076  
Db |||||  
QY 7915 ACGGTTTCAGATATTTCAGCTCCAGCATGGAGCCCTGTGTTCAACATCCGGTATGGGAC 7974  
QY 8077 AACGTTGGATGAGGAGAAAGGACCGGTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCCA 8136  
Db |||||  
QY 7975 AACTGTGAAAGAGGAAAGAAATCACGTGTTGGAGATTGCCAGACAGCGCGCATGGACA 8034  
QY 8137 AGCGTGGCCCGCAGAGCAGCAGAGACTGCGGGAAGGGGAGGAAGGCCCTGCGGGCCTGGAC 8196  
Db |||||  
QY 8035 GGCCTGGAATAAGGAACAAAGAGGCTGCAAGAGGGGGAAGAGGGGATTAGGGCATGGAC 8094  
QY 8197 AGAGGGGAGAAAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAGAGCTACGACGGCTTTT 8256  
Db |||||  
QY 8095 AGAAGGGGAAAAGCAGCAGCTTTTGAGCACTGGGCGGTACAAAGTTACGATGGGTATT 8154  
QY 8257 CGTGATCTCTGTCGAGCAGTACCCAGAACTGTGACAGCGCCAAACATCCACTTCAT 8316  
Db |||||  
QY 8155 TGTTTTGTCTGTTGAGCAGTATTAGAACCTTTCTGACAGTGCCTAATAATATTCACITTTAT 8214  
QY 8317 GAGACAGAGCGAGATGGGCCGAGGTGACAGAGA 8350  
Db |||||  
QY 8215 GAGACAGAGCGAAATAGGCAGGAGGTAACAAAA 8248

RESULT 14

ACC72052

ID ACC72052 standard; DNA; 9695 BP.

XX ACC72052;

XX 08-JUL-2003 (first entry)

XX BCU0205B gene #SEQ ID 81.

XX Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;  
KW drug discovery; clinical medicine; forensic medicine; gene;  
KW chromosome 5q33.3; ds.

XX Homo sapiens.

XX WO2003029421-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031287.

XX 03-OCT-2001; 2001US-0326526P.

XX 14-MAY-2002; 2002US-00144194.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;

XX WPI; 2003-381623/36.

XX P-PSDB; ABR58318.

XX New isolated human differentially-regulated breast cancer polynucleotide  
PT and polypeptide, useful for diagnosing, staging, prognosticating,  
PT preventing and/or treating diseases and conditions relating to breast  
PT cancer.

XX Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX The invention relates to isolated polynucleotides which are  
CC differentially-regulated in breast cancer. The methods and compositions  
CC of the present invention are useful for detecting, diagnosing, staging,  
CC monitoring, prognosticating, preventing and/or treating diseases and  
CC conditions relating to breast cancer, and may be used in gene therapy or



QY	2009	GAAGTGGACTGCATGGACCCACCATGTTTTCAGGCCGGGGTGTCTGCGTGAGAGGCGAATGC	2069
Db	2229	GAAGTTGATTGCTTGGATCCACCTGCTCCAGCCACGGAGTCTGTGTGAATGGAGAATGC	2288
QY	2069	CATTGCTTTGTGGGATGGGAGGCCACCAACTGCGGAGACCCCGAGGCCACATGCTTAGAC	2128
Db	2289	CTGTGCAGCCCTGGCTGGGTGGTCTGAACCTGTGAGTGGCGAGGTCCAGTGCCAGAC	2348
QY	2129	CAGTGTTCAGGCCACGGAACCTTCTCTCCCGGACACCGGGCTTTGCAGCTGTGACCCCAAGC	2188
Db	2349	CAGTGCACTGGGCATGGCAGCTACCTGCTTGACACGGGCCCTCTGCAGCTGCGATCCCAAC	2408
QY	2189	TGGACTGGACACGACTGTTTCTATCGAGATCTGTGCTGCCGACTGTGTTGGCCATGGCGTG	2248
Db	2409	TGGATGGGTCCCGACTGCTCTGTT-	2432
QY	2249	TGCGTAGGGGGCACCTGCTCCGCTGCGAGGATGGCTGGATGGGGGCAGCCTGCGACGAGCGG	2308
Db	2433	-	2432
QY	2309	GCCTGCCACCCGGCTGTGTGCCGAGCATGGGACCTGCCGCGACGGCAAGTGCAGATGCAGC	2368
Db	2433	-	2432
QY	2369	CCTGGCTGGAATGGCGAACACTGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAG	2428
Db	2433	-	2435
QY	2429	GGTTGCCCTGGGTTGTCAATGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCAC	2488
Db	2436	GGTGCCCTGACTTGTGCAACGGTAAACGGGAGATGCCACACTGGTTCAGAACAGCTGGCAG	2495
QY	2489	TGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACATTCATGGAGACTGCCTGC	2548
Db	2496	TGTGTCTGCCAGACCCGGCTGGAGAGGGCCCGGATGCAACGTTGCCATGGAACTTCCTGT	2555
QY	2549	GGTGACAGCAAGACAATGATGGAGATGGCCTGGTGGACTGCATGGACCCCTGACTGCTGC	2608
Db	2556	GCTGATAACAAGGATAATGAGGGAGATGGCCTGGTGGATGTTTGGACCCCTGACTGCTGC	2615
QY	2609	CTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCTAAACCCCTCTGGACATC	2668
Db	2616	CTGCAGTCAGCCTGTGAGAACAGCCTGCTCTGCCGGGGTCCCGGACCCACTGGACATC	2675
QY	2669	ATCCAGGAGACACAGTCCCTGTGTACAGCAGAACCTACACTCCTTCTATGACCGCATC	2728
Db	2676	ATTACG-----CAGGGCCAGACGGATTGGCCCGCAGTGAAGTCTTCTATGACCGTATC	2729
QY	2729	AAGTTCCTCGTGGGCAGGACAGCACGCACATAATCCCCGGGGAGAACCCCTTTGATGGA	2788
Db	2730	AAGCTCTTGGCAGGCAAGGATAGCACCCACATCATCTCTGGAGAGAACCCCTTCAACAGC	2789
QY	2789	GGGCATGCTTGTGTTATTGCTGGCCCAAGTGATGACATCAGATGGAACCCCTGCTGGTTGGT	2848
Db	2790	AGCTTGGTTTCTCTCATCCGAGGCCAAGTAGTAACCTACAGATGGAACCTCCCTGCTGGT	2849
QY	2849	GTGAACATCAGTTTGTCAATAACCTCTCTTTGGATATACAATCAGCAGGCAAGATGGC	2908
Db	2850	GTGAACGTGTCTTTGTCAAGTACCCAAATACGGGTACACCATCACCCCGCCAGGATGGC	2909
QY	2909	AGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCTCGGTTTCGAGCGGACACCT	2968
Db	2910	ACGTTTCGACCTGATCGCAAAATGGAGGTGCTTCCCTGACTCTACACTTTGAGCGAGCCCCG	2969
QY	2969	TTCATCACACAGGAGCACACCTGTGGCTGCCATGGGATCGCTTCTTTGTCTATGGAAACC	3028
Db	2970	TTCATGAGCCAGGAGCGCACTGTGTGGCTGCCGTGGAAACAGCTTTTACGCCATGGACACC	3029
QY	3029	ATCATCATGACATGAGGAGAATGAGATTCCCGACTGTGACCTGAGCAATTTTGGCCCGC	3088
Db	3030	CTGGTGATGAAGCCGAGGAGAACTCCATCCCCAGCTGTGACCTCAGTGGCTTTGTCCGG	3089
QY	3089	CCCAACCCAGTCTCTCTCCATCCCCACTGACGTCCTTCCCGAGCTCCTGTGCAGAGAAA	3148

Db	3090	CCTGATCCAATCATCATCTCTCTCCCACTGTCCACCTTCTTTAGTGTGCCCTTGGGCAG	3149
QY	3149	GGCCCCATTGTGCCGGAATTCAGGCTTTGAGGAGGAAATCTCTATCTCTGGTGCAG	3208
Db	3150	AATCCCATCGTGCCTGAGACCCAGGTTCTTCATGAAGAAATCGAGCTCCCTGGTTCCAAT	3209
QY	3209	ATGAGGCTGAGCTACCTGAGCAGCCGACCCCTGGCTACAAATCTGTCTCTGAGGATCAGC	3268
Db	3210	GTGAAACTTCGCTATCTGAGCTCTAGAACTGCAGGGTACAAGTCACTGCTGAAGATCACC	3269
QY	3269	CTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGGTAGCGTGGAG	3328
Db	3270	ATGACCCAGTCCACAGTGCCCTGAACCTCATAGGGTTACCTGATGGTGGTGTGCGAG	3329
QY	3329	GGCCGCCCTCTTCAGGAAGTGGTTCCGCTGCAGCCCCAGACCTGTCTCTATTATTTCAATTGG	3388
Db	3330	GGGCATCTCTTCCAGAAGTCATTCCAGGCTTCTCCAACTTGGCCTACACCTTCATCTGG	3389
QY	3389	GACAAGACAGACGCTCTACAACCAAGAGGTGTGGGCTTTCAGAAAGCCTTTGTTCCTGGT	3448
Db	3390	GACAAGACAGATGCCATATGGCCAAAGGGTGTATGGAATCTCAGATGCTGTTGTGTCTGTC	3449
QY	3449	GGTTATGAATATGAATCTCTGCCAGATCTAACTCTGTGGGAAAAAGAACAAACATCATGCC	3508
Db	3450	GGGTTTGAATATGAGACCTGTCCCAGTCTAATCTCTGGGAGAAAAGACAGCCCTCCTT	3509
QY	3509	CAGGGCTATGAAATTGACGCGTCCAAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCC	3568
Db	3510	CAGGGATTCCGAGCTGGACCCCTCCAACCTCGTGGCTGGTCCCTAGACAAACACACATC	3569
QY	3569	CTCAACATTCAAAGTGGTATCCTGCACAAAGGGAATGGGAGAACCAAGTTTGTGTCTCAG	3628
Db	3570	CTCAATGTTAAAGTGAATCCTACACAAAGSACCTGGGAAACCAAGTTCTGTGACCCAG	3629
QY	3629	CAGCCTCCTGTCTATTGGGAGCATCATGGGCAATGGGCGCCGGAGAGCATCTCCTGTCCCC	3688
Db	3630	CAGCCTGCCATCATCACCAGCATCATGGGCAATGGTGGCGCCGGAGCATTTCTGTGTCC	3689
QY	3689	AGTGCAACGGCCTTGCTGACGGCAACAAGCTCTGTGGCCCCAGTGGCCCCACCTGTGGC	3748
Db	3690	AGTGCAACGGCCTTGCTGAAGGCAACAAGCTGTGGCCCCAGTGGCTCTGGTGTGTGGA	3749
QY	3749	TCTGACGGGAGCCTCTATGTGGTGATTTCAACTACATTAGAAGGATCTTCCCTCTGTGA	3808
Db	3750	ATCGATGGGAGCCTCTATGTGGTGACTTCAATTAACATCGACGCATCTTTCCCTCTCGA	3809
QY	3809	AATGTCACCAACATCCTAGAGCTGAGGAATAAAGATTTAGACATAGTCACAGTCCAGCA	3868
Db	3810	AATGTGACCAGCATCTTGGAGTTACGAAATAAAGAGTTTAAACATAGCAACAACCCAGCA	3869
QY	3869	CACAAATACTACCTGGCCACAGACCCCATGAGTGGGCGCGTCTTCTCTCTGACAGCAAC	3928
Db	3870	CACAAGTACTACTTGGCAGTGGACCCCGTGTCCGCTCGCTCTACGTGTCCGACACCAAC	3929
QY	3929	AGCCGGCGGGTCTTTAAATCAAGTCCACTGTGGTGGTGAAGGACCTTGTCAAGAACTCT	3988
Db	3930	AGCAGGAGAATCTACCGCGTCAAGTCTCTGAGTGGAAACCAAGACCTGGCTGGGAATTCG	3989
QY	3989	GAGGTGTTGCGGGACAGGTGACCGAGTGCCTCCCTTTTGATGACACTCGCTCGCGGGAT	4048
Db	3990	GAAGTTGTGGCAGGACGGGAGAGCAGTGTCTACCTTTTGATGAAGCCCGCTGCGGGAT	4049
QY	4049	GGTGGGAAGGCCACAGAAGCCACACTCAACCAATCCAGGGGTATTACAGTGGACAAAGTTT	4108
Db	4050	GGAGGGAAGGCCATAGATGCAACCCCTGATGAGCCCGAGAGGTATTGAGTAGACAAGAAT	4109
QY	4109	GGGCTGATCTACTTCGTGGATGGCAACCATGATGACAGCGCATCGATCAGAATGGGATCATC	4168
Db	4110	GGGCTCATGTACTTTGTTCGATGCCACCATGATCCGGAAGGTTGACCAAGATGGAATCATC	4169
QY	4169	TCCACCCCTGCTCGGCTCTAATGATCTCACATCATCAGCCCGGCCACTCAGCTGTGATTTCTGTC	4228

Db 4170 TCCACCCCTGCTGGGCTCCAAATGACCTCACTGCGCGTCCGGCCGCTGAGCTGTGATTCCAGC 4229  
QY 4229 ATGGATATTTCCCAAGTAAGACTGGAGTGGCCACAGACTTAGCCATCAACCCAAATGGAC 4288  
Db 4230 ATGGATGTAGCCCAAGGTTGCTGAGTGGCCAAACAGACCTTGCTGTCAATCCCATGGAT 4289  
QY 4289 AACTCACTTTATGTCTCGACAACAAATGTGGTCTCTGCAAAATCTCTGAAAAACCCAGGTG 4348  
Db 4290 AACTCCTTGTATGTCTAGAGAACAAATGTCTTCCGAAATCACCGAGAACCCAAAGTC 4349  
QY 4349 CGCATTTGTCGGGAGGCCCATGCACTGCCAGGTCCCTGGCAATTGACCACTTCTTGCTA 4408  
Db 4350 AGCATCATTTGGGAGCGCCCATGCACTGCGCAAGTTCTTGGCAATTGACTACT--CACTC 4406  
QY 4409 AGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCCACCGCTTTGGCTGTTTTCACACAAT 4468  
Db 4407 AGCAAACTAGCCATTCACTCTGCCCTGGAGTCAGCCAGTGCCATTGCCATTCTCACACT 4466  
QY 4469 GGGGTCTCTGTATATGTCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCACC 4528  
Db 4467 GGGGTCTCTACATCACTGAGACAGATGAGAGAAAGATTAAACCGTCTACGCCAGGTAACA 4526  
QY 4529 ACTAGTGGAGAGATCTCACTCGTTGCTGGGCCCCCAGTGCGTGGCTGTAATAATAATGAT 4588  
Db 4527 ACCAACGGGAGATCTGCCCTTTTAGCTGGGCGAGCCTCGGACTGCGACTGCAAAAACGAT 4586  
QY 4589 GCCAACTGTGATTTTCTTGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACC 4648  
Db 4587 GTCAATTGCAACTGCTATTGAGAGATGATGCCCTACGCGACTGATGCCATCTTGAATTCC 4646  
QY 4649 CCATCTTCTTGGCTGTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAACATC 4708  
Db 4647 CCATCATCTTGTAGTCTCAGATCCAGATGTTACCATTTACATTGGAGACCTTGGAAATATT 4706  
QY 4709 CGAATTCGGTTTATCCGGAAGAACAAAGCCTTTCTCTCAACACCCAGAACATGTATGAGCTG 4768  
Db 4707 CGGATCAGGGCGGTGAGCAAGAACAAAGCCTGTTCTTAATGCTTCAACCATGATGAGGCT 4766  
QY 4769 TCTTCACCAATTGACCAGGAGCTCTATCTGTTTGATACCAACCGGCAAGCACCTGTACACC 4828  
Db 4767 GCATCCCCCGGAGACGAGGATTATATGTTTCAACGCTGATGGCATCCACCAATACACT 4826  
QY 4829 CAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACTACCTAGCTGGGACGGCGACATC 4888  
Db 4827 GTGAGCCTGGTGACAGGGGAGTACTTTGTIACAATTTTACATATAGTACTGACAATGATGTC 4886  
QY 4889 ACACATCATCACACAACAAATGGCAACATGGTAAATGTCCGCCGAGACTCTACTGGGATG 4948  
Db 4887 ACTGAATTGATTGACAATAATGGGAATTCCTGGAAGATCCGTCGGGACAGCATGGCATG 4946  
QY 4949 CCCCTCTGGCTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGT 5008  
Db 4947 CCCCCTCACCTGCTCATGCCCTGACAAACAGATCATCAACCTCACCGTGGGCACCAATGGA 5006  
QY 5009 GCACATCAAGAGTGTACACACAAGGACACGAGTTGGCCATGTATGACATACCATGGCAAT 5068  
Db 5007 GGCTTCAAAAGTGTGTCCACACAGAACCTGGAGCTTGGTCTCATGACCTATGATGGCAAC 5066  
QY 5069 TCCGGCCTTCTGGCAACCAAAAGCAATGAAAAAGGATGGACACATTTTATGAGTACGAC 5128  
Db 5067 ACTGGGCTCTTGGCCACCAAGAGCGATGAAACAGGATGGACGACTTCTATGACTATGAC 5126  
QY 5129 AGCTTTGGCCGCTGACAAATGTGACCTTCCCTACTGGCCAGGAGGAGCAGTTTCCGAAGT 5188  
Db 5127 CACGAAGGCCGCTGACCAACGCTGACGCGCCCCACGGGGTGTGTAACCAAGTCTGCAACCG 5186  
QY 5189 GATACAGACAGTTCACTGATGTCCAGGTPAGAGACCTCCA---GCAAGGATGATGTACAC 5245  
Db 5187 GAAATGGAGAAATCTATTACCATTTGACATTTGAGAACTCCAACCGTGTATGATGACGTCACT 5246  
QY 5246 ATAACCACCAACCTGTCTGCCTCAGGCGCTTCTACACACTGCTGCAAGACCAAGTCCGG 5305  
Db 5247 GTCATCACCAACCTCTCTTCAGTAGAGGCTCTCTACACAGTGGTACAAAGTCAAGTTCGG 5306

QY 5306 AACAGCTACTATATCGGGGCCGATGGCTCTCTTGGGCTGCTGCTGGCCAAACGGCATGGAG 5365  
Db 5307 AACAGCTACCAGCTCTGTAAATAATGTGTACCTGAGGTTGATGTATGTAATGGATGGT 5366  
QY 5366 GTGGCGCTGCAGACTGAGCCCCACTTGTGGCTGGCACCCGTCACCCACCGTGGGCAAG 5425  
Db 5367 ATCAGCTTCCACAGCGAGCCCCATGTCTAGCGGGCACCATCACCCACCATTTGGACGC 5426  
QY 5426 AGGAATGTACGCTGCCCATCGACAACCGCCTCAACCTGGTGGAGTGGCGCCAGCGCAA 5485  
Db 5427 TGCAACATCTCCCTGCTTATGGAGATGGCTTAAACTCCATTGAGTGGCGCTTAAGAAAG 5486  
QY 5486 GAGCAGGCTCGGGCCAGGTCACTGTCTTTGGCGCGGCTGCGGTCGACAAACCCGAAAT 5545  
Db 5487 GAACAGATTAAAGGCAAAAGTCACCATCTTTGGCAGGAAGCTCCGGTCCATGGAAGAAAT 5546  
QY 5546 CTCCTATCTCTGACTTTGATCGCGTAAACACGACAGAGAAAGATCTATGATGACCAACCGC 5605  
Db 5547 CTCTTGCTCACTGACTATGATCGAAATATTTCGACTGAAAGATCTATGATGACCAACCGG 5606  
QY 5606 AAGTTCACCCCTCGGATTCTGTACGACGAGCGGGCGGCCAGCCTCTGTCACCCAGC 5665  
Db 5607 AAGTTCACCCCTGAGGATCATTTATGACAGGTGGCGGCCCTTCTCTGGTGGCCAGC 5666  
QY 5666 AGCAGGCTGATGGTGTCAACGTGACATACTCCCTGGGGTTTACATTGCTGGCATCCAG 5725  
Db 5667 AGCGGCTGGCAGCTGTCAACGTGTCTACTTCTTCAATGGGGCGCTGGCTGGCTTCAG 5726  
QY 5726 AGGGGCATCATGTCTGAAAGAAATGGAATACGACAGCGGGCGCGCATCATCCAGGATC 5785  
Db 5727 CGTGGGCCATGAGCGAGAGGACAGACATCGACAAGCAAGCGCGCATCGTGTCCCGCATG 5786  
QY 5786 TTCGCTGATGGGAAGACATGGAGCTACACATACCTTAGAGAAATCCATGGTGTCTGCTACTA 5845  
Db 5787 TTCGCTGACGGGAAAGTGTGGAGTACTCTCTACCTTGACAAGTCCATGGTCTCTCTGCTT 5846  
QY 5846 CACAGCCAGAGGAGTATATCTTTGAGTTGACAAGAAATGACCGCCTCTCTTCTGTGACG 5905  
Db 5847 CAGAGCCAAACGTGAGTATATTTGAGTATGACTCCTCTGACCGCCTCCTTGCCGTCACC 5906  
QY 5906 ATGCCCAAACGTGGCGCGGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAAC 5965  
Db 5907 ATGCCCAGCGTGGCGCGGACAGCATGTCCACACACACACCTCCATCGGCTACATCCGTAAT 5966  
QY 5966 ATCTATCAGCCCTGAGGGCAATGCTCTCAGTACATACAGGACTTCACTGAGGATGGGCAC 6025  
Db 5967 ATTTACAACCCGCTGAAAGCAATGCTTCGGTCACTTTTGAAGTATGATGACGCGCCG 6026  
QY 6026 CTCCTTCAACACTTCTACCTGGGCACTGGCGGAGGTGATATACAAGTATGGCAAACTG 6085  
Db 6027 ATCCTGAAGACCTCTTTTGGCACCGGACGCGGAGTGTCTTACAAGTATGGAAACTC 6086  
QY 6086 TCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTGAGTTTCACTTATGACGAGACG 6145  
Db 6087 TCCAAGTTATCAGAGATTGTCTACGACAGTACCGCGCTCACCTTCGGGTATGACGAGACC 6146  
QY 6146 GCAGGCATGTGAAGACCATCAACCTACAGAATGAGGGCTTCACTGCACCATCCGCTAC 6205  
Db 6147 ACTGGTGTCTTGAAGATGGTCAACCTCCAAAGTGGGGCTTCTCTGCACCATCAGGTAC 6206  
QY 6206 CGTCAGATTGGCCCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTC 6265  
Db 6207 CGGAAGATTGGCCCCCTGGTGGACACAGCATCTACAGGTTCTCCGAGGAAGGCATGGTC 6266  
QY 6266 AACGCCGTTTGTACTACAACT---ATGACAACAGCTTCCGGGTGACCAAGCATGCAGGCT 6322  
Db 6267 AATGCCAGGTTTGACTACACCTTATCATGACAACAGCTTCCGATCGCAAGCATCAAGCCC 6326  
QY 6323 GTGATCAACGAGACCCCACTGCCCAATTGATCTCTATCGTATGATGATGTGTGAGGCAAG 6382  
Db 6327 GTCATAAGTGAGACTCCCTCCCTCCCGGTGACCTCTACCGCTATGATGATGATTTCTTGGCAAG 6386

Qy	6383	ACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACCAAGATCATCAACCACA	6444
Db	6387	GTGGAACACTTTGGTAAAGTTTGGAGTCACTATTATGACATCAACCAGATCATCAACCACT	6446
Qy	6443	GCTGTCAATGACCAACACCAAGCAATTTTGATGCATATGGCAGGATGAAGGAAGTGCAGTAT	6502
Db	6447	GCCGTGATGACCTCAGCAAAACACTTCGACACCCATGGCGGATCAAGGAGTCCAGTAT	6506
Qy	6503	GAGATCTTCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACAATGGGCGAGTA	6562
Db	6507	GAGATGTTCCGCTCCCTCATGTACTGGATGACCGTCAATATGACAGCATGGCAGGGTG	6566
Qy	6563	GTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGTACTCCTATGAGTAT	6622
Db	6567	ATCAAGAGGGAGCTAAAACCTGGGGCCCTATGCCAATACCAAGTACACCTATGACTAC	6626
Qy	6623	GATGCTGACGSSCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGC	6682
Db	6627	GATGGGACGSGCAGCTCCAGAGCGTGGCCGTCAATGACCGCCGACCTGGCGCTACAGC	6686
Qy	6683	TACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGAAACAGTGCACGGCTCACACCA	6742
Db	6687	TATGACCTTAAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGCGCCTCATGCC	6746
Qy	6743	CTACGGTATGACATCCGCGACCGCATCACTCGGTGGTGACGTGCAATACAGATGGAT	6802
Db	6747	TTGCGCTATGACCTCCGGGATCGGATAACAGACTCGGGGATGTGCAGTACAAAATTGAC	6806
Qy	6803	GAGGATGGCTTCTGAGGCAGCGGGCGGTGATATCTTTTGAGTACAACTCAGCTGGCCTG	6862
Db	6807	GACGATGGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAAATTCACACTGAC	6866
Qy	6863	CTCATCAAGGCCTACAACCGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCTG	6922
Db	6867	CTAACAGAGCCTACAAACAGGCCAGCGGGTGAGTGTCCAGTACCGCTATGATGGCGTA	6926
Qy	6923	GGCGGGCGGTGTCAGCAAGAGCAGCCACAGCAACCACTGCAGTCTTCTATGACAGAC	6982
Db	6927	GGACGGCGGCTTCCTACAAGACCAACCTGGGCCACCACTGCAGTACTTCTACTCTGAC	6986
Qy	6983	CTGACCAACCCCAAGGTCAACCCACCTGTACAACCACTCCAGCTCTGAGATCACCTCC	7042
Db	6987	CTCCACAACCCGACGCGCATCACCCATGTCTACAATCACTCCAACCTCGGAGATTACCTCA	7046
Qy	7043	CTCTACTACGACTTGAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGAAGTAT	7102
Db	7047	CTGTACTACGACCTCCAGGGCCACCTCTTTGCCATGGAGAGCAGCAGTGGGAGGAGTAC	7106
Qy	7103	TACATAGCTTGTGACAACATCGGGACCCCTCTTGTGTCTTTAGTGAACAGGTTTGATG	7162
Db	7107	TATGTTGCCTCTGATAACACAGGGACTCCTCTGGCTGTGTTTCAGCATCAACGGCCTCATG	7166
Qy	7163	ATCAAGCAAAATCCTGTACACAGCCCTATGGGGAGATCTACATGGATACCAACCCCAACTTT	7222
Db	7167	ATCAAACAGCTGCAGTACACGGCCTATGGGGAGATTATTATGACTCCAACCCCGACTTC	7226
Qy	7223	CAGATCATCATAGGCTACCATGGTGGCCCTCTATGATCCACTCACCAAGCTTGTCACATG	7282
Db	7227	CAGATGGTTCATTGGCTTCCATGGGGACTCTATGACCCCTTGACCAAGCTGGTCCACTTC	7286
Qy	7283	GGCCGGCGAGATTATGATGTGCTGGCCGACGCTGGACTAGCCAGACCAAGCTGTGG	7342
Db	7287	ACTCAGCGTGATTATGATGTGCTGGCAGGACGATGGACCTCCCCAGACTATACCATGTGG	7346
Qy	7343	AAGCACCTTAGTAGCAGCAACGTCACTGCGCTTTTAATCTCTATATGTTCAAAAACAACAAC	7402
Db	7347	AAAAACGTGGCAAGGAGCCGGCC--CCCTTTAACTGTATATGTTCAAGAGCAACAAT	7403
Qy	7403	CCCATCAGCAACTCCAGGACATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGCTC	7462
Db	7404	CCTCTCAGCAGTGAGCTAGATTTGAAGAACTACGTGACAGATGTGAAAAGCTGGCTGTG	7463
Qy	7463	ACCTTTGGATCCAGCTACACAACGTGATCCCTGGTTATCCCAAAACAGACATGGATGCC	7522

Db	7464	ATGTTTGGATTTTCAGCTTAGCAACATCATCTCTGGCTTCCCAGAGAGCCAAATGTATTTC	7523
QY	7523	ATGGAACCCCTCCTACGAGCTCATCCACACACAGATGAAAACGCAGGAGTGGGACAAACAGC	7582
Db	7524	GTGCCTCTCCCTATGA-----ATTGTCAGAGAGTCAAGCAAGTGAATGGA	7571
QY	7583	AAGTCATCCTCGGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCCCTTTGTACACCTTA	7642
Db	7572	CAGCTCATTAACAGGTGTCCAAACAGACACAGAGAGACATAACCCAGGCCCTTCATGGCTCTG	7631
QY	7643	GAACGGTTTGACCAAGCTCTATGGCTCCAAATCACCAAGTGCACAGGCTCCAAAGACC	7702
Db	7632	GAA-----GGACAGGTCACTTACTATAAAAGCTCCACGCCAGCATCCGAGAGAAAAGCAGGT	7685
QY	7703	AAGAAAGTTTGCAATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTGAAGGAT	7762
Db	7686	CAGTGGTTTGCCACCAACCAACCGCCCATCATTTGGCAAGGCATCATGTTTGCCATCAAAGAA	7745
QY	7763	GGCCGAGTGACCAACAGACATCATCAGTGTGGCCCAATGAGGATGGGCGAAGGTTGCTGCC	7822
Db	7746	GGCGGGTGACCAACGGGCGTGTCCAGCATGCGCCAGCGAAGATAGCCGAAGGTGGCATCT	7805
QY	7823	ATCTTGAAACATGCCCACTACCTAGAGAACCTGCATTCACCATTTGATGGGTTGGATACC	7882
Db	7806	GTGCTGAACAACGCCTACTACCTGGACAAGATGCATACAGCATCGAGGGCAAGGACACC	7865
QY	7883	CATTACTTTGTGAAAACCAAGGACCTTCAGAAGGTGACCTGGCCATCTCTGGCCCTCAGTGGG	7942
Db	7866	CACACTTTGTGAAGATTGGCTCAGCCGATGGCGACCTGGTCACACTAGGCACCAACATC	7925
QY	7943	GGCGGGCGAAACCTTGAGAAATGGGTCACCTCACTGTGTCCCAGATCAACACACAGTACTT	8002
Db	7926	GGCGCGAAGGTGCTAGAGAGCGGGGTGAACGTGACCGTGTCCCAGCCACGCTGCTGGTC	7985
QY	8003	AATGSCAGGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAAC	8062
Db	7986	AACGGCAGGACTCGAAGTTTCACGAACATTGAGTTCAGTACTCCACGCTGCTGCTCAGC	8045
QY	8063	ACACGCTACGG-----GACAACTTTGGATGAGGAGAAAGGCACGGTCTCTGGAGCTG	8113
Db	8046	ATCCGCTATGGCTTCACCCCGACACCCCTGGACGAAGAGAAAGCCCGCTCTCTGGACCAG	8105
QY	8114	GCCCGGCAGAGACCGGTGCGCCCAAGCTGGSCCCCGAGCAGCAGAGACTCGGSGAAGGG	8173
Db	8106	CGGAGACAGAGGGCCCTGGGCACGGCCTGGGCCAAGGAGCAGCAGAAAGCCAGGGAACGGG	8165
QY	8174	GAGGAAGCCTCGGGCCCTGGACAGAGGGGGAGAAAGCAGCAGGTGCTGAGCACAGGGCCGG	8233
Db	8166	AGAGAGGGAGCCGCTGTGGACTGAGGGCCGAGAAAGCAGCAGCTTCTGAGCACCGGGCGC	8225
QY	8234	GTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTTCAGAC	8293
Db	8226	GTGCAAGGTTACGAGGGATATTACGTGCTTCCCGTGGAGCAATACCCAGAGCTTCAGAC	8285
QY	8294	AGCGCCAAACAACATCCACTTTCATGAGACAGACGAGATGGCCCGAGGTGACAGA	8348
Db	8286	AGTAGCAGCAACATCCAGTTTTTTAAGACAGAATGAGATGGAAAGAGGTAAACAA	8340
RESULT 15			
ID	ACC72051	standard; DNA; 9058 BP.	
XX	ACC72051;		
AC	ACC72051;		
XX	08-JUL-2003	(first entry)	
XX	BCU0205A	gene #SEQ ID 79.	
DE	BCU0205A	gene #SEQ ID 79.	
XX	Breast cancer;	cytostatic; gene therapy; antisense therapy; regulated;	
KW	drug discovery;	clinical medicine; forensic medicine; gene;	
KW	chromosome 5q33.3;	ds.	

RESULT 15  
ACC72051  
ID ACC72051 standard; DNA; 9058 BP.  
XX  
AC ACC72051;  
XX  
DT 08-JUL-2003 (first entry)  
XX  
DE BCU0205A gene #SEQ ID 79.  
XX  
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated  
KW drug discovery; clinical medicine; forensic medicine; gene;  
KW chromosome 5q33.3; ds.







Db 6118 TGACCTCCGGATCGGATAACACAGACTCGGGATGTGCAGTACAAAATTGACGACGATGG 6177  
QY 6811 CTTTCCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATCAA 6870  
Db 6178 CTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATAAATTCGAAGGGCCTCCTAAACAG 6237  
QY 6871 GGGCTACAACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCCTGGGGCGGCG 6930  
Db 6238 AGCTTACAACAAGGCCAGCGGGTGGAGTGTCCAGTACCGCTATGATGGCTAGGACGGCG 6297  
QY 6931 CGTGTCCAGCAAGAGCAGCAACAGCCAGCCACCCTGAGTCTGAGATCACCTCCCTCTACTA 7050  
Db 6298 GGGTTCCTACAAGACCAACCTGGGCCACCACCTGCGAGTACTTCTACTCTGACCTCCACAA 6357  
QY 6991 CCCACCAAGGTCAACCCACCTGTACAAACCACTCCAGCTCTGAGATCACCTCCCTCTACTA 7050  
Db 6358 CCCAGCGGCATCAACCCATGCTTACAATCACTCCAACTCGGAGATTACCTCACTGTACTA 6417  
QY 7051 CGACTTGCAAGGACACCTCTTTGCCATGGAGCTGAGAGTGGTGTGATGATTTTACATAGC 7110  
Db 6418 CGACTCCAGGGCCACCTCTTTGCCATGGAGAGCAGCAGTGGGAGGAGTACTATGTTGC 6477  
QY 7111 TTGTGACAAACATCGGGACCCCTCTTGTCTCTTTAGTGAACAGGTTTGTGATGATCAAGCA 7170  
Db 6478 CTCTGATAACACAGGGACTCTCTGGCTGTGTTTCAGATCAACGGCCTCATGATCAACAA 6537  
QY 7171 AATCCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCAACCTTCAGATCAT 7230  
Db 6538 GCTGAGTACACGGCCCTATGGGAGATTTATTATGACTCCAACCCCGACTTCAGATGGT 6597  
QY 7231 CATAGGCTACCATGGTGGCCTCTATGATCCACTCAACAAGCTTGTCCACATGGGCCGGCG 7290  
Db 6598 CATGGCTTCCATGGGGACTCTATGACCCCTTGACCCAGCTGGTCCACTTCACTCAGCG 6657  
QY 7291 AGATTATGATGTGCTGGCCGACGCTGGACTAGCCAGACCAAGCTGTGAGACACCT 7350  
Db 6658 TGATTATGATGTGCTGGCAGGACGATGGACCTCCCAAGACTATACCATGTGGAACAAAGT 6717  
QY 7351 TAGTAGCAGCAACGTATGCTCTTTTAATCTCTATATGTTCAAAAACAAACCCCATCAG 7410  
Db 6718 ---GGGCAAGGAGCGCGCCCTTTAACTGTATATGTTCAAGAGCAACAATCCTCTCAG 6774  
QY 7411 CAACCTCCAGGACATCAAGTGTCTCATGACAGATGTTAAAGCTGGCTGCTCACCTTTGG 7470  
Db 6775 CAGTGAGTAGATTGGAAGAACTACGTGACAGATGTGAAAGCTGGCTTGTGATGTTGG 6834  
QY 7471 ATTCCAGCTACACAACGTGATCCCTGGTTATCCCAAACCAAGACATGGATGCCATGGAACC 7530  
Db 6835 ATTTACGCTTAGCAACATCATTCCTGGCTTCCCGAGAGCCAAAATGTATTTCTGTCCTCC 6894  
QY 7531 CTCCTACGAGCTCATCCACACACAGATGAAAACGAGGAGTGGGACAAACAGCAAGTCTAT 7590  
Db 6895 TCCCTATGAATT-----GTCAAGAGTCAAGCAAGTGAGAAATGGACAGCTCAT 6942  
QY 7591 CCTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGTCACTTTAGAACGGTT 7650  
Db 6943 TACAGGTGTCCAACAGACAACAGAGAGACATAACCAAGGCCCTTCATGGCTCTGGAAGGA-- 7000  
QY 7651 TGACGAGCTCTATGGCTCCACAATCACCAGCTGCCAGCGAGTCCCAAAGACCAAGAGTT 7710  
Db 7001 ----CAGGTCAATTAAAGCTCCAGCCAGCATCCGAGAGAAAGCAGGTCACTGGTT 7056  
QY 7711 TGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGT 7770  
Db 7057 TGCCACCACACGCCCATCATTTGGCAAGGCATCATGTTTGCATCAAGAAAGGGCGGT 7116  
QY 7771 GACCACAGACATCATCAGTGTGGCCAAATGAGGATGGCGGAAGGTTGTGCCATCTTCAA 7830  
Db 7117 GACCACGGGGGTGTCCAGCATCGCCAGCGAAGATAGCCGCAAGGTGGCATCTGTGCTCAA 7176  
QY 7831 CCATGCCCACTACCTAGAGAACCTTGCACTTCAACATTGATGGGGTGGATACCCATTACTT 7890

Db 7177 CAACGCTACTACTCTGGACAAGATGCACTACAGCATCGAGGGCAAGGACACCCACTACTT 7236  
QY 7891 TGTGAAAACAGGACCTTTCAGAAAGTGCACCTGGCCATCCTGGGCTCCTCAGTGGGGCGGCG 7950  
Db 7237 TGTGAAGATTGGCTCAGCCGATGGCGACCTGGTCACACTAGGCACCACCATCGGCCGCAA 7296  
QY 7951 AACCTGGAGAAATGGGTCAACGTCACTGTGTCTCCAGATCAACACAGTACTTTAATGGCAG 8010  
Db 7297 GGTGTAGAGAGCGGGTGAACGTGACCGTGTCCAGCCACGCTGCTGTCAACCGGAG 7356  
QY 8011 GACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAACACACGCTA 8070  
Db 7357 GACTCGAAGGTTTACGAACAATTGAGTTCAGTACTCCACGCTGCTGCTCAGCATCCGCTA 7416  
QY 8071 CGG-----GACAAACGTTTGATGAGGAGAAAGSCACGGGTCTTGAGCTGGCCCGCA 8121  
Db 7417 TGGCCTCACCCCGACACCCCTGGACGAAGAAAGSCCGCTCCTGGACCAAGCGGAGACA 7476  
QY 8122 GAGAGCGCTCGCCCAAGCGTGGGCCCGGAGCAGCAGACTGCGGGAAGGGGAGGAAGG 8181  
Db 7477 GAGGGCCTGGGCACCGGCTGGCCAAAGGAGCAGCAGAAAGCCAGGACCGGAGAGAGG 7536  
QY 8182 CCTGCGGCTTGACAGAGGGGAGAAAGCAGCAGGTGCTGAGCACAGGGCGGTGCAAGG 8241  
Db 7537 GAGCGCCTGTGGACTGAGGGCGGAGAAAGCAGCAGCTTCTGAGCACCGGGCGCTGCAAG 7596  
QY 8242 CTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTCAAGACAGCGCCAA 8301  
Db 7597 GTACGAGGATATTACGTGCTTCCCTGGAGCAATACCCAGAGCTTGCAGACAGTAGCAG 7656  
QY 8302 CAACATCCACTTTCATGAGACAGACGAGATGGGCCGAGGTGACAGA 8348  
Db 7657 CAACATCCAGTTTAAAGACAGATGAGATGGAAAGAGGTAAACAAA 7703

Search completed: June 25, 2004, 00:48:24  
Job time : 2061 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 17:15:34 ; Search time 366 Seconds  
(without alignments)  
12666.830 Million cell updates/sec

Title: US-10-029-020-13  
Perfect score: 8354  
Sequence: 1 gtttggatgtggaggagc.....ccggaggtgacagagaggac 8354

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1708.8	20.5	3111	3	US-08-891-845-1 Sequence 1, Appli
2	1708.8	20.5	3111	4	US-09-514-573-1 Sequence 1, Appli
3	1673.6	20.0	1680	3	US-08-891-845-3 Sequence 3, Appli
4	1673.6	20.0	1680	4	US-09-514-573-3 Sequence 3, Appli
5	1185.6	14.2	2387	3	US-08-891-845-11 Sequence 11, Appl
6	1185.6	14.2	2387	4	US-09-514-573-11 Sequence 11, Appl
7	389	4.7	447	4	US-09-833-381-1072 Sequence 1072, Ap
8	208.6	2.5	2007	4	US-09-976-594-407 Sequence 407, App
9	109	1.3	13857	4	US-09-620-312D-75 Sequence 75, Appl
10	92.2	1.1	6049	4	US-08-793-273C-3 Sequence 3, Appli
11	92.2	1.1	6049	5	PCT-US95-11684-3 Sequence 3, Appli
12	84	1.0	7286	4	US-08-793-273C-1 Sequence 1, Appli
13	84	1.0	7286	5	PCT-US95-11684-1 Sequence 1, Appli
14	58.4	0.7	7218	1	US-08-232-463-14 Sequence 14, Appl
15	57	0.7	4724	1	US-08-404-665-3 Sequence 3, Appli
16	57	0.7	4724	1	US-08-404-671-3 Sequence 3, Appli
17	57	0.7	4724	1	US-08-404-781-3 Sequence 3, Appli
18	56	0.7	4403765	3	US-09-103-840A-2 Sequence 2, Appli
19	56	0.7	4411529	3	US-09-103-840A-1 Sequence 1, Appli
20	54.2	0.6	4411529	3	US-09-103-840A-1 Sequence 1, Appli
21	52.6	0.6	4403765	3	US-09-103-840A-2 Sequence 2, Appli
22	49.8	0.6	690	4	US-09-252-991A-3288 Sequence 3288, Ap
23	49.8	0.6	1173	4	US-09-252-991A-3422 Sequence 3422, Ap
24	49.8	0.6	1434	4	US-09-252-991A-3260 Sequence 3260, Ap
25	49.6	0.6	293	4	US-09-313-294A-6760 Sequence 6760, Ap
26	49.6	0.6	3489	2	US-08-728-323A-1 Sequence 1, Appli
27	49.6	0.6	3489	4	US-09-298-568-1 Sequence 1, Appli

c	28	49.6	0.6	3489	4	US-09-410-399-1	Sequence 1, Appli
	29	49.6	0.6	32207	2	US-08-770-379-20	Sequence 20, Appl
	30	49.6	0.6	32207	3	US-08-757-669A-20	Sequence 20, Appl
	31	49.6	0.6	32207	4	US-09-230-371A-20	Sequence 20, Appl
	32	48.2	0.6	2595	4	US-09-810-268-2	Sequence 2, Appli
	33	48.2	0.6	3007	4	US-09-810-268-1	Sequence 1, Appli
c	34	47.2	0.6	319	3	US-09-165-264-8	Sequence 8, Appli
	35	47.2	0.6	1863	4	US-09-252-991A-733	Sequence 733, App
	36	47.2	0.6	1980	4	US-09-252-991A-762	Sequence 762, App
	37	47	0.6	152331	3	US-09-128-155-16	Sequence 16, Appl
c	38	46.2	0.6	762	4	US-09-252-991A-781	Sequence 781, App
	39	46.2	0.6	1278	4	US-09-252-991A-615	Sequence 615, App
	40	46.2	0.6	1404	4	US-09-252-991A-636	Sequence 636, App
	41	45.4	0.5	1439	4	US-09-056-556-167	Sequence 167, App
	42	45.4	0.5	1439	4	US-09-072-596-162	Sequence 162, App
	43	45.4	0.5	1439	4	US-09-072-967-167	Sequence 167, App
	44	45.2	0.5	2064	1	US-08-343-428-1	Sequence 1, Appli
	45	45.2	0.5	6530	2	US-08-146-930-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-891-845-1  
; Sequence 1, Application US/08891845  
; Patent No. 6096873  
; GENERAL INFORMATION:  
; APPLICANT: Schaefer, Gabriele M.  
; APPLICANT: Sliwowski, Mark  
; TITLE OF INVENTION: Gamma-Herregulin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/891,845  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/021640  
; FILING DATE: 07/12/96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P1043  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3111 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
US-08-891-845-1

Query Match 20.5%; Score 1708.8; DB 3; Length 3111;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1713; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTTTGTGGATGTGGAGGAGCGCGCGGAGCCATGGACGTGAAGGAGGAGCCCTTA 60  
|||||

Db	300	GTTTGTGGATGTGGAGGAGCGGGGCCGGAGGCCATTGGACGTGAAGGAGAGGAAGCCTT	359
QY	61	CGCTCGCTGACCCGGCGCGGACGCCGAGCGCGCTACACAGACTCGTCCGCGGACAG	120
Db	360	CGCTCGCTGACCCGGCGCGGACGCCGAGCGCGCTACACAGACTCGTCCGCGGACAG	419
QY	121	CGAGGAGGCAAAAGCCCGCAGAAATCGTACAGCTCCAGCAGACCCCTGAAGCCTACGA	180
Db	420	CGAGGAGGCAAAAGCCCGCAGAAATCGTACAGCTCCAGCAGACCCCTGAAGCCTACGA	479
QY	181	CCAGACGCCCGCCTAGCCTATGGCAGCCGCGTCAAGGACATTGTCCGCAGGAGGCCGA	240
Db	480	CCAGACGCCCGCCTAGCCTATGGCAGCCGCGTCAAGGACATTGTCCGCAGGAGGCCGA	539
QY	241	GGAATTCTGCCGCACAGGTGCCAACTTCACCCCTCGGGAGCTGGGGCTGGAAGAAAGTAAC	300
Db	540	GGAATTCTGCCGCACAGGTGCCAACTTCACCCCTCGGGAGCTGGGGCTGGAAGAAAGTAAC	599
QY	301	GCCCCCTCACGGGACCCCTGTACCCGACAGACATTGGCCCTGCCCTACTCCCAT	360
Db	600	GCCCCCTCACGGGACCCCTGTACCCGACAGACATTGGCCCTGCCCTACTCCCAT	659
QY	361	GGGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTCCCTGAGCACCCCGT	420
Db	660	GGGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTCCCTGAGCACCCCGT	719
QY	421	GCGTCTGTGGGGCCGGACACACGGTCAGGGCGAGCTCTGCCTGTCCAGCCGGGCCAA	480
Db	720	GCGTCTGTGGGGCCGGACACACGGTCAGGGCGAGCTCTGCCTGTCCAGCCGGGCCAA	779
QY	481	TTCCAAATCTCACACTCACCGACACCGAGCATGAAACAATGAGACTGATCATCCGGCGG	540
Db	780	TTCCAAATCTCACACTCACCGACACCGAGCATGAAACAATGAGACTGATCATCCGGCGG	839
QY	541	CCTGCAGAACCAACGGCGGCTCCGGACGCCCGCGCGCTCTCGACCGCCACACCCC	600
Db	840	CCTGCAGAACCAACGGCGGCTCCGGACGCCCGCGCGCTCTCGACCGCCACACCCC	899
QY	601	CAACACAGCACACCGGGCTCCATTAACTCCCTGAACCGGGCAACTTCACGCCGAGGAG	660
Db	900	CAACACAGCACACCGGGCTCCATTAACTCCCTGAACCGGGCAACTTCACGCCGAGGAG	959
QY	661	CAACCCAGCCCGGCCCCACGACCACTCGCTCTCCGAGAGCCCCCTGCCGGCGGCG	720
Db	960	CAACCCAGCCCGGCCCCACGACCACTCGCTCTCCGAGAGCCCCCTGCCGGCGGCG	1019
QY	721	CCAGGAGCCTGCCACGCCCAGGAGAACTGGTGTCTCAACAGCAACATCCCCCTGGAGAC	780
Db	1020	CCAGGAGCCTGCCACGCCCAGGAGAACTGGTGTCTCAACAGCAACATCCCCCTGGAGAC	1079
QY	781	CAGGAACCTAGGCAAGCAGCCATTCTTAGGGACATTGCAGGACAACTCAATTGAGATGA	840
Db	1080	CAGGAACCTAGGCAAGCAGCCATTCTTAGGGACATTGCAGGACAACTCAATTGAGATGA	1139
QY	841	CATTCTCGCGCCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTCCTCTTCAAGCC	900
Db	1140	CATTCTCGCGCCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTCCTCTTCAAGCC	1199
QY	901	TGGAGGCACCTCCCGCTCTTCTGCACCACATCACAGGGTACCCACTGACGTCCAGCAC	960
Db	1200	TGGAGGCACCTCCCGCTCTTCTGCACCACATCACAGGGTACCCACTGACGTCCAGCAC	1259
QY	961	AGTGTAATCTCTCCCGCCCGACCCCTGCCCGCAGCACCTTCGCCCGCGGCGCTTTAA	1020
Db	1260	AGTGTAATCTCTCCCGCCCGACCCCTGCCCGCAGCACCTTCGCCCGCGGCGCTTTAA	1319
QY	1021	CCTCAAGAAAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCCATCGTCAT	1080
Db	1320	CCTCAAGAAAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCCATCGTCAT	1379
QY	1081	CTCAGCCACTCTGGTCATCTCTGCTGGCATACTTTGTGGCCATGCACCTGTTTGGCCTAAA	1140
Db	1380	CTCAGCCACTCTGGTCATCTCTGCTGGCATACTTTGTGGCCATGCACCTGTTTGGCCTAAA	1439

RESULT 2  
US-09-514-573-1  
; Sequence 1, Application US/09514573  
; Patent No. 6500941  
; GENERAL INFORMATION:  
; APPLICANT: Schaefer, Gabriele M.  
; APPLICANT: Sliwkowski, Mark  
; TITLE OF INVENTION: Gamma-Heregulin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/514,573  
; FILING DATE: 28 FEB 2000  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/891845  
; FILING DATE: 10 JULY 1997  
; APPLICATION NUMBER: 60/021640  
; FILING DATE: 07/12/96  
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
;
US-09-514-573-1

Query Match      20.5%; Score 1708.8; DB 4; Length 3111;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 GTTTGTGGATGTGGAGGAGCGCGGGCCGAGGCCATGGACGCTGAAGGAGAGGAAGCCTTA 60
Db      1 |||||
QY      300 GTTTGTGGATGTGGAGGAGCGCGGGCCGAGGCCATGGACGCTGAAGGAGAGGAAGCCTTA 359
Db      300 |||||
QY      61 CCGCTCGCTGACCCCGGCGCCCGGACGCCGCGCTACACGACTCGTCCCGGACAG 120
Db      61 |||||
QY      360 CCGCTCGCTGACCCCGGCGCCCGGACGCCGCGCTACACGACTCGTCCCGGACAG 419
Db      360 |||||
QY      121 CGAGGAGGGCAAGCCCCCGCAGAAATCGTACAGTCCAGCGAGACCCCTGAAGGCTACGA 180
Db      121 |||||
QY      420 CGAGGAGGGCAAGCCCCCGCAGAAATCGTACAGTCCAGCGAGACCCCTGAAGGCTACGA 479
Db      420 |||||
QY      181 CCAGGACGCCCGCTAGCCTATGGCAGCCGCGTCAAGGACATTGTGCCGACGAGGCCCGA 240
Db      181 |||||
QY      480 CCAGGACGCCCGCTAGCCTATGGCAGCCGCGTCAAGGACATTGTGCCGACGAGGCCCGA 539
Db      480 |||||
QY      241 GGAATTCTGCCGACAGGTGCCAACTTCACCCCTGGGGAGCTGGGGCTGGAAGAAGTAAC 300
Db      241 |||||
QY      540 GGAATTCTGCCGACAGGTGCCAACTTCACCCCTGGGGAGCTGGGGCTGGAAGAAGTAAC 599
Db      540 |||||
QY      301 GCCCCCTCACGGGACCCTGTACCGGACAGACATTGGCCCTGCCCAATGCGGCTACTCCAT 360
Db      301 |||||
QY      600 GCCCCCTCACGGGACCCTGTACCGGACAGACATTGGCCCTCCCCCACTGCGGCTACTCCAT 659
Db      600 |||||
QY      361 GGGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTGCCCTGAGCACCCCGT 420
Db      361 |||||
QY      660 GGGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGTCTGCCCTGAGCACCCCGT 719
Db      660 |||||
QY      421 GCGTCTGTGGGCGGGAGCACACGGTCAAGGGCGGAGCTCTGCTGCTCCCTGAGCACCCGCAA 480
Db      421 |||||
QY      720 GCGTCTGTGGGCGGGAGCACACGGTCAAGGGCGGAGCTCTGCTGCTCCAGCCGGGCGCAA 779
Db      720 |||||
QY      481 TTCCAATCTCACACTCACCGGACACCGAGCATGAATAACACTGAGACTGATCATCCGGGCGG 540
Db      481 |||||
QY      780 TTCCAATCTCACACTCACCGGACACCGAGCATGAATAACACTGAGACTGATCATCCGGGCGG 839
Db      780 |||||
QY      541 CCTGCAGAACCAACGCGGGCTCCGACGCCCGCGCGCTCTCGCACGCCCAACACCCC 600
Db      541 |||||
QY      840 CCTGCAGAACCAACGCGGGCTCCGACGCCCGCGCGCTCTCGCACGCCCAACACCCC 899
Db      840 |||||
QY      601 CAACCCAGCACCAACGCGGCTCCATTAACTCCCTGAACCGGGGCAACTTCACGCCGAGGAG 660
Db      601 |||||
QY      900 CAACCCAGCACCAACGCGGCTCCATTAACTCCCTGAACCGGGGCAACTTCACGCCGAGGAG 959
Db      900 |||||
QY      661 CAACCCAGCCCGGCCCGCCACGGACCACTCGCTCTCCGAGAGCCCCCTGCGGGCGGCGC 720
Db      661 |||||
QY      960 CAACCCAGCCCGGCCCGCCACGGACCACTCGCTCTCCGAGAGCCCCCTGCGGGCGGCGC 1019
Db      960 |||||
QY      721 CCAGGAGCCTGCCACGCCCCAGGAGAACTGGTGTCTCAACAGCAACATCCCCCTGGAGAC 780
Db      721 |||||
QY      1020 CCAGGAGCCTGCCACGCCCCAGGAGAACTGGTGTCTCAACAGCAACATCCCCCTGGAGAC 1079
Db      1020 |||||
QY      781 CAGGAACCTAGGCAAGCAGCCATTTCCTAGGGACATTGCAGGACAACTCATTTGAGATGGA 840
Db      781 |||||

```



Qy	1655	TTGGCTTTTACAATGACGGAAGGAGTCAGAAGTGGTTTCTTTCTCACCACGCCATT	1714
Db	1621	TTGGCTTTTACAATGACGGAAGGAGTCAGAAGTGGTTTCTTTCTCACCACGCCATT	1680

RESULT 4  
US-09-514-573-3  
; Sequence 3, Application US/09514573  
; Patent No. 6500941  
; GENERAL INFORMATION:  
; APPLICANT: Schaefer, Gabriele M.  
; APPLICANT: Sliwowski, Mark  
; TITLE OF INVENTION: Gamma-Heregulin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/514,573
FILING DATE: 28 FEB 2000
CLASSIFICATION:

```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/891845  
FILING DATE: 10 JULY 1997  
APPLICATION NUMBER: 60/021640  
FILING DATE: 07/12/96  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P1043  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1680 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
US-09-514-573-3

Query Match 20.0%; Score 1673.6; DB 4; Length 1680;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	35	ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCCGCGACGCCGAGCGC	94
Db	1	ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCCGCGACGCCGAGCGC	60
QY	95	CGCTACACCCAGCTCGTCCCGGGACAGCGAGGAGGGCAAAAGCCCGCAGAAATCGTACAGC	154
Db	61	CGCTACACCCAGCTCGTCCCGGGACAGCGAGGAGGGCAAAAGCCCGCAGAAATCGTACAGC	120
QY	155	TCCAGCGGAGACCTGAAGGCCTACGACCCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC	214
Db	121	TCCAGCGGAGACCTGAAGGCCTACGACCCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC	180
QY	215	AAGGACATTGTGCCGACGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCAACCCCTG	274
Db	181	AAGGACATTGTGCCGACGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCAACCCCTG	240
QY	275	CGGGAGCTGGGGTGGGAAGAAGTAACGCCCCCTCAGCGGACCCCTGTACCCGGACAGACATT	334

QY 1415 CCTGTGCATCTGAAATTC AATGTGTCTCTGGGAAAAGGCAGCCCTGGTTGGCAATTTATGGC 1474  
Db 1381 CCTGTGCATCTGAAATTC AATGTGTCTCTGGGAAAAGGCAGCCCTGGTTGGCAATTTATGGC 1440  
QY 1475 AGAAAAGGCTCCCTCCTTCA CATACACAGTTTGACTTTTGTGGAGCTGCTGGATGGCAGG 1534  
Db 1441 AGAAAAGGCTCCCTCCTTCA CATACACAGTTTGACTTTTGTGGAGCTGCTGGATGGCAGG 1500  
QY 1535 AGGCTCCTAACCCAGGAGGCGGAGCCTAGAGGGGACCCCGGCCAGTCTCGGGGAACT 1594  
Db 1501 AGGCTCCTAACCCAGGAGGCGGAGCCTAGAGGGGACCCCGGCCAGTCTCGGGGAACT 1560  
QY 1595 GTGCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCAC 1654  
Db 1561 GTGCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCAC 1620  
QY 1655 TTGGCTTTTACATGACCGAAAGGAGTCAGAAAGTGTTTCTTCTCACCCTGCTGCAAT 1714  
Db 1621 TTGGCTTTTACATGACCGAAAGGAGTCAGAAAGTGTTTCTTCTCACCCTGCTGCAAT 1680

RESULT 5

US-08-891-845-11  
; Sequence 11, Application US/08891845  
; Patent No. 6096873  
; GENERAL INFORMATION:  
; APPLICANT: Schaefer, Gabriele M.  
; APPLICANT: Sliwowski, Mark  
; TITLE OF INVENTION: Gamma-Herregulin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/891,845  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/021640  
; FILING DATE: 07/12/96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P1043  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2387 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
US-08-891-845-11

Query Match 14.2%; Score 1185.6; DB 3; Length 2387;  
Best Local Similarity 98.8%; Pred. No. 1.5e-290;  
Matches 1194; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 530 CATCCGGGGCGGCTGCAGAACCCAGCGGGCTCCGGACGCCCGCGCGCTCTCGCAC 589  
Db 1 CATCCGGGGCGGCTGCAGAACCCAGCGGGCTCCGGACGCCCGCGCGCTCTCGCAC 60

QY 590 GCCACACACCCCAACACAGCACACACGGGCGCTCCATTAACCTCCCTGAACCGGGGCAACTTC 649  
Db 61 GCCACACACCCCAACACAGCACACACGGGCGCTCCATTAACCTCCCTGAACCGGGGCAACTTC 120  
QY 650 ACGCCGAGGAGCAACCCAGCCCGGCCCCCACGGACCACTCGCTCTCCGAGAGCCCCCT 709  
Db 121 ACGCCGAGGAGCAACCCAGCCCGGCCCCCACGGACCACTCGCTCTCCGAGAGCCCCCT 180  
QY 710 GCCGGCGGCGCCAGGAGCCTGCCACGCCCCAGGAGAACTGGTGCTCAACAGCAACATC 769  
Db 181 GCCGGCGGCGCCAGGAGCCTGCCACGCCCCAGGAGAACTGGTGCTCAACAGCAACATC 240  
QY 770 CCCCTGGAGACAGGAACCTAGGCAAGCAGCCATTTCCTAGGACATTCAGGACAACTC 829  
Db 241 CCCCTGGAGACAGGAACCTAGGCAAGCAGCCATTTCCTAGGACATTCAGGACAACTC 300  
QY 830 ATTGAGATGGACATTTCTCGGCGCTCCCGCCCATGATGSSGCTTACAGTGACGGGCACITC 889  
Db 301 ATTGAGATGGACATTTCTCGGCGCTCCCGCCCATGATGSSGCTTACAGTGACGGGCACITC 360  
QY 890 CTCTTCAAGCCTGGAGGCACCTCCCGCGCTCTTCTGCACCACATCACAGGTTACCCACTG 949  
Db 361 CTCTTCAAGCCTGGAGGCACCTCCCGCGCTCTTCTGCACCACATCACAGGTTACCCACTG 420  
QY 950 ACGTCCAGCAGTGTACTCTCTCTCGGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGG 1009  
Db 421 ACGTCCAGCAGTGTACTCTCTCTCGGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGG 480  
QY 1010 CCGGCTTTAACTCAAGAAAGCCCTCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 1069  
Db 481 CCGGCTTTAACTCAAGAAAGCCCTCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540  
QY 1070 GCCATCGTCACTCAGCCACTCTGTGTCATCTCTGTGGCAATCTTGTGGCCATGCACCTG 1129  
Db 541 GCCATCGTCACTCAGCCACTCTGTGTCATCTCTGTGGCAATCTTGTGGCCATGCACCTG 600  
QY 1130 TTTGGCCTAAACTGGCACCTTGCCGATGGAGGGCAGATGTATGAGATCACGGAGGAC 1189  
Db 601 TTTGGCCTAAACTGGCACCTTGCCGATGGAGGGCAGATGTATGAGATCACGGAGGAC 660  
QY 1190 ACAGCCAGCAGTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGCACTGGC 1249  
Db 661 ACAGCCAGCAGTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGCACTGGC 720  
QY 1250 TTAGAGACCCCTGACAGGAAGGCAAGGAACCAAGAAAGCCAGTAGTTTCTTT 1309  
Db 721 TTAGAGACCCCTGACAGGAAGGCAAGGAACCAAGAAAGCCAGTAGTTTCTTT 780  
QY 1310 CCAGAGGACAGTTTCTAGATTTCTGGAGAAATTTGATGTGGAAAGGCGAGCTCCAGAG 1369  
Db 781 CCAGAGGACAGTTTCTAGATTTCTGGAGAAATTTGATGTGGAAAGGCGAGCTTCCAGAG 840  
QY 1370 ATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGAAA 1429  
Db 841 ATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGAAA 900  
QY 1430 TTCAATGTGTCTCTGGGAAAGGCGACCCCTGGTTGGCAATTTATGGCAGAAAAGGCCCT 1489  
Db 901 TTCAATGTGTCTCTGGGAAAGGCGACCCCTGGTTGGCAATTTATGGCAGAAAAGGCCCT 960  
QY 1490 CCTTCACATACACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAG 1549  
Db 961 CCTTCACATACACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAG 1020  
QY 1550 GAGGCGGAGCCTTAGAGGGGACCCCGCGCCAGTCTCGGGGAACTGTGCCCCCTCCAGC 1609  
Db 1021 GAGGCGGAGCCTTAGAGGGGACCCCGCGCCAGTCTCGGGGAACTGTGCCCCCTCCAGC 1080  
QY 1610 CATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTTGGCTTTTACAT 1669  
Db 1081 CATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTTGGCTTTTACAT 1140

Qy	1670	GACGGAAAGGAGTCAGAAAGTGGTTTCCTTTCTCACCACTGCCATTGAGTCGGTGGATAAC	1729
Db	1141	GACGGAAAGGAGTCAGAAAGTGGTTTCCTTTCTCACCACTGCCATTGATTCAGTGGTACA	1200
Qy	1730	TGCCCCAG	1737
Db	1201	GGACAGAG	1208

## RESULT 6

US-09-514-573-11  
; Sequence 11, Application US/09514573  
; Patent No. 6500941  
; GENERAL INFORMATION:  
; APPLICANT: Schaefer, Gabriele M.  
; APPLICANT: Slikowski, Mark  
; TITLE OF INVENTION: Gamma-Heregulin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb fl  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/514,573  
; FILING DATE: 28 FEB 2000  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/891845  
; FILING DATE: 10 JULY 1997  
; APPLICATION NUMBER: 60/021640  
; FILING DATE: 07/12/96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P1043  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2387 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
US-09-514-573-11

Db	181	GCCGGCGGGCCACAGGAGCCTGCCACGCGCCAGGAGAACTGGCTGCTCAACAGCAACATC	240
QY	770	CCCCTGGAGACCAAGAAACCTAGGCAAGCAGSCATTCTCTAGGGACATTGCAGGACAAACCTC	829
Db	241	CCCCCTGGAGACCAAGAAACCTAGGCAAGCAGSCATTCTCTAGGGACATTGCAGGACAAACCTC	300
QY	830	ATTGAGATGSAATTTCTCGGCGCCTCCCGCCATGATGGGCTTTACAGTACGCGGCACTTC	889
Db	301	ATTGAGATGSAATTTCTCGGCGCCTCCCGCCATGATGGGCTTTACAGTACGCGGCACTTC	360
QY	890	CTCTTTCAAGCCTGGAGGCACCTCCCGCTCTTTCTGCACCACATCACCCAGSGTACCCACTG	949
Db	361	CTCTTTCAAGCCTGGAGGCACCTCCCGCTCTTTCTGCACCACATCACCCAGSGTACCCACTG	420
QY	950	ACGTCCAGCACAGTGTACTCTCTCTCCGCCCCGACCCCTGCCCGCGCAGCACCTTCGCCCGG	1009
Db	421	ACGTCCAGCACAGTGTACTCTCTCTCCGCCCCGACCCCTGCCCGCGCAGCACCTTCGCCCGG	480
QY	1010	CCGGCCTTTAACTTCAAGAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTTGAGC	1069
Db	481	CCGGCCTTTAACTTCAAGAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTTGAGC	540
QY	1070	GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTTGTGGCCATGCACCTG	1129
Db	541	GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTTGTGGCCATGCACCTG	600
QY	1130	TTTGGCCCTTAACTGGCACCTTGCAGCCGATGGAGGGGCAGATGTATGAGATCACGGAGGAC	1189
Db	601	TTTGGCCCTTAACTGGCACCTTGCAGCCGATGGAGGGGCAGATGTATGAGATCACGGAGGAC	660
QY	1190	ACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCCTATACCCCTCAGGGGGCACTGGC	1249
Db	661	ACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCCTATACCCCTCAGGGGGCACTGGC	720
QY	1250	TTAGAGACCCCTGACAGGAAAGGCAAGAAACACAGAAAGGAAAGCCACAGTAGTTTCTTT	1309
Db	721	TTAGAGACCCCTGACAGGAAAGGCAAGAAACACAGAAAGGAAAGCCACAGTAGTTTCTTT	780
QY	1310	CCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGAAAGGCGAGCTCCACAGAAG	1369
Db	781	CCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGAAAGGCGAGCTCCACAGAAG	840
QY	1370	ATTCCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAA	1429
Db	841	ATTCCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAA	900
QY	1430	TTCAATGTGTCTCTGGAAAGGCAGCCCTGGTTGGCATTTATGGCAGAAAAGGCCCTCCCT	1489
Db	901	TTCAATGTGTCTCTGGAAAGGCAGCCCTGGTTGGCATTTATGGCAGAAAAGGCCCTCCCT	960
QY	1490	CCTTCACATACACAGTTTGTGACTTTGTGAGCTGCTGGATGGCAGGAGSGTCTTAACCCAG	1549
Db	961	CCTTCACATACACAGTTTGTGACTTTGTGAGCTGCTGGATGGCAGGAGSGTCTTAACCCAG	1020
QY	1550	GAGGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGTGCCCCCCTCCAGC	1609
Db	1021	GAGGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGTGCCCCCCTCCAGC	1080
QY	1610	CATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCATTGGCTTTTACAAT	1669
Db	1081	CATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCATTGGCTTTTACAAT	1140
QY	1670	GACGGAAGGAGTCAGAAGTGGTTTCTTCTCACCACCTGCCATTGAGTCGGTGGATAAC	1729
Db	1141	GACGGAAGGAGTCAGAAGTGGTTTCTTCTCACCACCTGCCATTGATCCAGTGGTACA	1200
QY	1730	TGCCCCAG	1737
Db	1201	GGACAGAG	1208

RESULT 7  
US-09-833-381-1072

```
; Sequence 1072, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1072
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(447)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1072

Query Match      4.7%; Score 389; DB 4; Length 447;
Best Local Similarity 94.0%; Pred. No. 8e-89;
Matches 426; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 2341 CTGCCGCGACGGGAAGTGGAGTGCAGCCCTGGCTGGAAATGGCAACACTGCACCATCGC 2400
Db 22 CTGCCGCGACGGGAAGTGGAGTGCAGCCCTGGCTGGAAATGGCAACACTGCACCATC-- 79

QY 2401 TCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGCTTGTGCAATGGCAACGGCAG 2460
Db 80 -----GAGGGTTGCCCTGGCTTGTGCAATGGCAACGGCAG 114

QY 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGCTGG 2520
Db 115 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGCTGG 174

QY 2521 CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAGACAATGATGGAGATGGCCT 2580
Db 175 CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAGACAATGATGGAGATGGCCT 234

QY 2581 GGTGGACTGCATGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640
Db 235 GGTGGACTGCATGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 294

QY 2641 CCTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGAGACACAGTCCCTGTGTACAGCA 2700
Db 295 CCTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGAGACACAGTCCCTGTGTACAGCA 354

QY 2701 GAACCTACACTCCTTCTATGACCCGCATCAAGTTCTCGTGGGAGGACAGCACCGCACAT 2760
Db 355 GAACCTACACTCCTTCTATGACCCGCATCAAGTTCTCGTGGGAGGACAGCACCGCACAT 414

QY 2761 AATCCCCGGGAGAACCCCTTTTGATGGAGGCA 2793
Db 415 AATCCCCGGGAGAACCCCTTTTGATGGAGGCA 447

RESULT 8
US-09-976-594-407/c
; Sequence 407, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
```

```
; SOFTWARE: PERL Program
; SEQ ID NO 407
; LENGTH: 2007
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 474200.1
US-09-976-594-407

Query Match      2.5%; Score 208.6; DB 4; Length 2007;
Best Local Similarity 72.7%; Pred. No. 1.1e-42;
Matches 296; Conservative 0; Mismatches 109; Indels 2; Gaps 2;

QY 7941 GGGGGCGGCGAACCCCTGGAGAAATGGGTCAACGTCACCTGTGTCCAGATCAACACAGTAC 8000
Db 2007 GCGGGCGCAAGCGCTGGAGAACGGCATCAACGTCACGCTGTGTCCAGTCCACACCGGTGG 1948

QY 8001 TTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACT-GTGCTTG 8059
Db 1947 TGATCGGCAGGACGCGCAGGTTCCGCGACGTGGAGATGCAGTTCGGCGCTGGCGCTGC 1888

QY 8060 AACACACGCTACGGGACAAACGTTGGATG-AGGAGAAGGCACGGGTCTCTGGAGCTGGCCCG 8118
Db 1887 CACGTGGCTACGGCATGACCCCTGGACGAAGGAGAGGCGCGCATCTCTGGAGCAGCGCG 1828

QY 8119 GCAGAGAGCCCTGCGCCCAAGCGTGGGCGCGGAGCAGCAGAGACTGCGGGGAAGGGAGGA 8178
Db 1827 GCAGCGCGCGCTCGCCCGCGCTGGCGCGGAGCAGCAGCGCGTGGCGCAGCGCGAGGA 1768

QY 8179 AGGCCTGCGGCGCTGGACAGAGGGGAGAGAGCAGAGTGTGAGCAGACAGGGCGGTGCA 8238
Db 1767 GGGCGCGCGCTCTGGACGGAGGGGAGAGCGGAGCTGTGAGCGCGCGCAAGGTGCA 1708

QY 8239 AGGTACGACGCGCTTTTTCGTGATCTCTGTGAGCAGTACCAGAACTGTGAGACAGCGC 8298
Db 1707 GGGCTACGACGCGTACTACGTACTCTCGTGGAGCAGTACCCGAGCTGGCCGACAGCGC 1648

QY 8299 CAACAACATCCACTTCATGAGACAGAGCCAGATGGGCCGAGGTGAC 8345
Db 1647 CAACAACATCCAGTTCTCTCGGCGAGAGCGAGATCGGCAGGAGGTAAC 1601

RESULT 9
US-09-620-312D-75
; Sequence 75, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
```





Db 895 CCTCTGTGCTCAACAAT-----TGCTACAACCGTGGACGATGCGTGGAGAAATGAG 945  
QY 1970 TGCATCTGCAACCCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCC 2029  
Db 946 TGCCTGTGTGATGAGGGTTTACGGGCGAAGACTGCGAGTGAGCTCATCTGCCCAATGAC 1005  
QY 2030 ACATGTTTCAAGCCGGGTGTCTGCGTGTGAGAGGCGAATGCCAATTGCTTTTGGATGGGA 2089  
Db 1006 ---TGCTTCGACCGGGCGCGCTGCATCAATGGCACCTGCTACTGCGAAGAAGCTTCACA 1062  
QY 2090 GGCACCAACTGCGAGACCCCGAGGCCACATGCTTTAGACCAGTGTTCAGGCCACGGAACC 2149  
Db 1063 GGTGAAGACTGCGGGAACCCACCTGCCCA-----CATGCCTGCCACACCCAG 1110  
QY 2150 TTCTCTCCGGACACCGGGCTTTGCAGCTGTGACCCAAAGCTGGACTGGACACGACTGTTCT 2209  
Db 1111 GGCGGTGTGAGGAGGGCAGTGTATGTATGATGAGGCTTTGCCGGTGTGGACTGCAGC 1170  
QY 2210 ATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGCACCTGCCGC 2269  
Db 1171 GAGAAGAGGTGTCCTGCTGACTGTCACAATCGTGGCGCTGTGTAGACGGGCGGTGTGAG 1230  
QY 2270 TGCAGGATGGCTGGATGGGGCAGCCTGCGACCGAGCGGCTGCCACCCGCGCTGTGCC 2329  
Db 1231 TGTGATGATGTTTCACTGGAGCTGACTGTGGGAGCTCAAGTGTCCCAATGCTGCAGT 1290  
QY 2330 GAGCATGGGACCTGCCCGCAGCGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAACAC 2389  
Db 1291 GGCCATGGCCGCTGTGTCAATGGGCAGTGTGTGTGATGAGGGCTATATCTGGGAGGAC 1350  
QY 2390 TGCACC 2395  
Db 1351 TGCAGC 1356

RESULT 13

PCT-US95-11684-1  
; Sequence 1, Application PC/TUS9511684  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE  
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING  
; TITLE OF INVENTION: AND USING SAME  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 North Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/11684  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/308,359  
; FILING DATE: 16-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April C.  
; REGISTRATION NUMBER: 33,950  
; REFERENCE/DOCKET NUMBER: BEC0019P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; TELEFAX: 619-554-6312  
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7286 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 55..6654  
; OTHER INFORMATION: /product= "cytotactin"  
PCT-US95-11684-1  
  
Query Match 1.0%; Score 84; DB 5; Length 7286;  
Best Local Similarity 49.1%; Pred. No. 1e-10;  
Matches 327; Conservative 0; Mismatches 315; Indels 24; Gaps 3;  
  
QY 1730 TGCCCCAGCAACTGCTATGGCAATGGTACTGCTATCTCTGGACCTGCCACTGCTTCTCTG 1789  
Db 715 TGCCCCAGCGACTGCAATGACCAGGCAAGTGCCTGATGGAGTCTGCATCTGTTTCGAA 774  
QY 1790 GGTTCCTGGGCCCCGACTGTGGCAGAGCCCTCCTGCCCGTCTCTGTAGCGGAATGGC 1849  
Db 775 GGCTACCGGCTGACTGCAGCCGTGAATCTGCCAGTGCCTGCAGTGAGGAGCAGGC 834  
QY 1850 CAATACATGAAAGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGCTGAGTGCATGTG 1909  
Db 835 ACATGTGTAGATGGCTTGTGTGTGCCACGATGGCTTTCAGGCGGATGACTGCAACAAG 894  
QY 1910 CCCACCAACAGTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCAACGGGACC 1969  
Db 895 CCTCTGTGTCTCAACAAT-----TGCTACAACCGTGGACGATGCGTGGAGAAATGAG 945  
QY 1970 TGCATCTGCAACCCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCC 2029  
Db 946 TGCCTGTGTGATGAGGGTTTACGGGCGAAGACTGCAGTGAAGTCACTGCCCCAATGAC 1005  
QY 2030 ACATGTTTCAAGCCGGGTGTCTGCGTGTGAGAGGCGAATGCCAATTGCTTTTGGGATGGGA 2089  
Db 1006 ---TGCTTCGACCGGGCGCGCTGCATCAATGGCACCTGCTACTGCGAAGAAGGCTTCACA 1062  
QY 2090 GGCACCAACTGCGAGACCCCGAGGCCACATGCTTTAGACCAGTGTTCAGGCCACGGAACC 2149  
Db 1063 GGTGAAGACTGCGGGAACCCACCTGCCCA-----CATGCCTGCCACACCCAG 1110  
QY 2150 TTCTCTCCGGACACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTCT 2209  
Db 1111 GGCGGTGTGAGGAGGGCAGTGTGTATGTATGATGAGGGCTTTGCCGGTGTGGACTGCAGC 1170  
QY 2210 ATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCGC 2269  
Db 1171 GAGAAGAGGTGTCTCTGCTGACTGTCAATCGTGGCCGCTGTGTAGACGGGCGGTGTGAG 1230  
QY 2270 TGCAGGATGGCTGGATGGGGGCGAGCCTCGACCCAGCGGGCTTCCACCCCGCTGTGCC 2329  
Db 1231 TGTGATGATGGTTTCACTGGAGCTGACTGTGGGAGCTCAAGTGTCCCAATGGCTGCAGT 1290  
QY 2330 GAGCATGGGACCTGCCCGCGACCGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACAC 2389  
Db 1291 GGCCATGGCCGCTGTGTCAATGGGCAGTGTGTGTGATGAGGGCTATATCTGGGAGGAC 1350  
QY 2390 TGCACC 2395  
Db 1351 TGCAGC 1356

RESULT 14

US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.

```

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

```

```

Query Match 0.7%; Score 58.4; DB 1; Length 7218;
Best Local Similarity 4.4%; Pred. No. 0.00031;
Matches 17; Conservative 220; Mismatches 151; Indels 0; Gaps 0;

QY 7894 GAAACCAGGACCTTCAGAGGTGACCTGGCCATCTGGGCTCAGTGGGGGGGCGGGAAC 7953
Db 1427 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1368

QY 7954 CTGGAGAAATGGGGTCAACGTCACTGTGCCAGATCAACACAGTACTTAATGGCAGAC 8013
Db 1367 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1308

QY 8014 TAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAACACACGCTACGG 8073
Db 1307 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1248

QY 8074 GACAAAGTTGGATGAGGAGAGGACGGGCTCTGGAGCTGGCCCGGAGAGAGCGCTGG 8133
Db 1247 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1188

QY 8134 CCAAGCGTGGCCCGGAGCAGCAGAGACTCGGGAAGGGAGGAGGCGCTGGGCGCTG 8193
Db 1187 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1128

QY 8194 GACAGAGGGGAGAGCAGCAGAGGTGCTGAGCAGAGGGCGGGTCAAGGCTACGACGGCTT 8253
Db 1127 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1068

QY 8254 TTTCTGTGATCTCTGTCGAGCAGTACCCA 8281
Db 1067 RATCGAAGCTCCCTCGACCTGCAGCCA 1040

```

```

RESULT 15
US-08-404-665-3
; Sequence 3, Application US/08404665
; Patent No. 5591583
; GENERAL INFORMATION:
; APPLICANT: Reid, Robert A.
; APPLICANT: Ackley, Rhonda L.
; APPLICANT: Hemperly, John J.
; TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404,665
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3341
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4724 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-404-665-3

```

```

Query Match 0.7%; Score 57; DB 1; Length 4724;
Best Local Similarity 50.9%; Pred. No. 0.00055;
Matches 135; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 2129 CAGTGTTCAGCCACGGAACCTTCTCCCGGACACCGGGCTTTGACAGCTGTGACCCAAAGC 2188
Db 985 CACTGCAGTGGCCACGGCAACTTTAGCTTTGAGTCTGTGGTCTGATCTGCAACGAAGGC 1044

QY 2189 TGGACTGGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTG 2248
Db 1045 TGGTTTGGCAAGAATTGCTCGGAGCCCTACTGCCCGCTGGTGTGCTCCAGCCGGGGGTG 1104

QY 2249 TGCGTAGGGGACCTGCCGCTGCGAGGATGGCTGGATGGGGGAGCCTGCGACCGCGG 2308
Db 1105 TGTGTGATGCGCCAGTGCATCTGTGACAGCGAGTACAGCGGGGATGACTGTTCCGAATC 1164

QY 2309 GCCTGCCACCCGCTGTGCCGAGCATGGGACCTGCCCGGACCGCAAGTGCAGTGCAGC 2368
Db 1165 CGGTGCCCAACAGACTGCAGCTCCCGGGGCTCTGCGTGGACGGGAGTGTGTCTGTGAA 1224

QY 2369 CCTGGCTGGAATGGGAACTGCA 2393
Db 1225 GAGCCCTACACTGGCGAGGACTGCA 1249

```

Search completed: June 25, 2004, 00:13:59  
Job time : 383 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 17:14:23 ; Search time 18786 Seconds  
(without alignments)  
13279.498 Million cell updates/sec

Title: US-10-029-020-13

Perfect score: 8354

Sequence: 1 gtttgtggtgaggagc.....ccggaggtgacagaggac 8354

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: em\_gss\_hum:\*
- 18: em\_gss\_inv:\*
- 19: em\_gss\_pln:\*
- 20: em\_gss\_vrt:\*
- 21: em\_gss\_fun:\*
- 22: em\_gss\_mam:\*
- 23: em\_gss\_mus:\*
- 24: em\_gss\_pro:\*
- 25: em\_gss\_rod:\*
- 26: em\_gss\_phg:\*
- 27: em\_gss\_vrl:\*
- 28: gb\_gss1:\*
- 29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5857.4	70.1	6246	29	AY413475 Homo sapi
2	5045.8	60.4	5970	29	AY413476 Pan trogl
3	2353	28.2	5094	29	AY405420 Homo sapi
4	2347.8	28.1	3190	29	AY413477 Mus muscu

5	2282.2	27.3	5069	29	AY405421 Pan trogl
6	2232.8	26.7	5087	29	AY405422 Mus muscu
7	1407.8	16.9	3038	11	AK037897 Mus muscu
8	743.8	8.9	1284	11	AK039472 Mus muscu
9	713.8	8.5	718	9	AL045768 DKFZp434F
10	695.2	8.3	856	14	CF534506 UI-M-GIO-
11	630.8	7.6	827	12	BM947883 UI-M-EGOp
12	629	7.5	870	10	BG036207 602326960
13	613.6	7.3	728	14	CF536185 UI-M-GIO-
14	609.2	7.3	770	14	CB524534 UI-M-FY0-
15	602.6	7.2	2627	11	AK050784 Mus muscu
16	581	7.0	1031	29	CNS02PHB Tetraodon
17	580.8	7.0	728	14	CB520016 UI-M-GIO-
18	567.2	6.8	572	13	BX504178 DKFZp686D
19	561.4	6.7	668	13	BQ563239 gi01b09.Y
20	544.8	6.5	694	14	CF723353 UI-M-GV0-
21	534.6	6.4	637	14	CF171639 B0845C10-
22	531.2	6.4	798	10	BF144774 601791454
23	528.6	6.3	632	10	BB655038 BB655038
24	525	6.3	621	14	CF182790 UI-M-EY0-
25	518.6	6.2	525	9	AL046228 DKFZp434E
26	514	6.2	592	10	BF349559 M30-HT015
27	497	5.9	693	13	BQ210132 UI-R-EP0-
28	494.2	5.9	620	14	CB532612 757081 MA
29	491	5.9	620	14	CB532245 756697 MA
30	489.4	5.9	566	13	BU072782 im47g02.Y
31	486	5.8	900	13	BU267121 603506284
32	485.8	5.8	894	13	BU151768 AGENCOURT
33	473.2	5.7	545	10	BF075317 224264 MA
34	473	5.7	734	13	BU233488 603408464
35	471.4	5.6	675	10	BE374675 601224176
36	469.2	5.6	603	10	BB611598 BB611598
37	468.6	5.6	574	9	AV589581 AV589581
38	464	5.6	600	14	CA527167 8032-72 M
39	456.2	5.5	868	29	CNS03W3H Tetraodon
40	455.2	5.4	668	13	BU214946 603106244
41	451.6	5.4	724	13	BU387957 603583120
42	449.8	5.4	819	13	BQ443891 UI-M-EW0-
43	449.8	5.4	951	13	BU839812 AGENCOURT
44	446.8	5.3	757	10	BE741784 601595546
45	443.4	5.3	461	10	BF932388 IL2-NT019

ALIGNMENTS

RESULT 1

AY413475

LOCUS

DEFINITION

AY413475 Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES Location/Qualifiers  
source 1..6246  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
gene <1..>6246  
/locus\_tag="HCM4903"  
ORIGIN  
Query Match 70.1%; Score 5857.4; DB 29; Length 6246;  
Best Local Similarity 95.4%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 233; Indels 57; Gaps 4;  
QY 2055 TGAGAGCGGAATGCCATGCTTTGTGGGATGGGGAGGCACCAACTGGGAGACCCCGAGGG 2114  
DB 2 TGAAGGCAGATGCTGTGCCACAGTGGCTGGAAGGCGCTGAGTGGGATGTGCCACCA 61  
QY 2115 CCACATGCTTAGACCAAGTTCAGGCCACGGAACCTTCTCCCGGACACCGGCTTTGCA 2174  
DB 62 ACCAGTGATCGATGCGCTGCAGCAACCATGGCACCTGC---ATCACGGGCACCTGCA 118  
QY 2175 GCTGTGACCCCAAGCTGGACTGGACACGACTGTTCTATCGAGATCTGTGCTGCGACTGTG 2234  
DB 119 TCTGCAACCTGGCTACAAGGGCGAGAGCTGTGAGGAAGAGATCTGTGCTGCGACTGTG 178  
QY 2235 GTGGCCATGGCGTGCCTAGGGGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGGCGAG 2294  
DB 179 GTGGCCATGGCGTGCCTAGGGGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGGCGAG 238  
QY 2295 CCTGCGACCAAGCGGCTGCCACCCCGCGCTGTGCCGAGCATGGGACCTGCCGCGACGGCA 2354  
DB 239 CCTGCGACCAAGCGGCTGCCACCCCGCGCTGTGCCGAGCATGGGACCTGCCGCGACGGCA 298  
QY 2355 AGTGGAGTGCAGCCCTGGCTGGAATGGGGAACACTGCACCATCGCTCACTATCTGGATA 2414  
DB 299 AGTGGAGTGCAGCCCTGGCTGGAATGGGGAACACTGCACCATC----- 342  
QY 2415 GGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGACC 2474  
DB 343 -----GNNN 391  
QY 2475 TGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCA 2534  
DB 392 NNN 451  
QY 2535 TGGAGACTGCGCTGCGGTGACAGCAAGCAATGATGGAGATGGGCTGGTGGACTGCTGG 2594  
DB 452 NNN 511  
QY 2595 ACCCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCGCTA 2654  
DB 512 ACCCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCGCTA 571  
QY 2655 ACCCTCTGGACATCATCCAGGAGACAGAGTCCCTGTGTCAAGCAGAGAACCTACACTCCT 2714  
DB 572 ACCCTCTGGACATCATCCAGGAGACAGAGTCCCTGTGTCAAGCAGAGAACCTACACTCCT 631  
QY 2715 TCTATGACCGCATCAAGTTCTCGTGGGAGGACAGCAGCAGCATATAATCCCGGGGAGA 2774  
DB 632 TCTATGACCGCATCAAGTTCTCGTGGGAGGAGCAGCAGCAGCATATAATCCCGGGGAGA 691  
QY 2775 ACCCCTTTGATGAGGGGCATGCTTGTGTTATTCGTGGCCAAAGTATGACATCAGATGGA 2834  
DB 692 ACCCCTTTGATGAGGGGCATGCTTGTGTTATTCGTGGCCAAAGTATGACATCAGATGGA 751  
QY 2835 CCCCCTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATATACATCA 2894  
DB 752 CCCCCTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATATACATCA 811

2895 GCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGGGCATCTCATCATCTCCTGCGGT 2954  
DB 812 GCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGGGCATCTCATCATCTCCTGCGGT 871  
QY 2955 TCGAGCGGGCACCTTTTCATCACACAGGAGCACACCTGTGGCTGCCATGGGATCGCTTCT 3014  
DB 872 TCGAGCGGGCACCTTTTCATCACACAGGAGCACACCTGTGGCTGCCATGGGATCGCTTCT 931  
QY 3015 TTGTTCATGGAACCATCATCATGAGACATGAGGAGAAATTCAGGCTTTGCAGGAGGAAATCTCTA 3074  
DB 932 TTGTTCATGGAACCATCATCATGAGACATGAGGAGAAATTCAGGCTTTGCAGGAGGAAATCTCTA 3194  
QY 3135 CCTGTGCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCTTTGCAGGAGGAAATCTCTA 3194  
DB 1052 CCTGTGCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCTTTGCAGGAGGAAATCTCTA 1111  
QY 3195 TCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCCTGGCTACAAATCTG 3254  
DB 1112 TCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGACCCCTGGCTACAAATCTG 1171  
QY 3255 TCCTGAGGATCAGCCTCACCCACCCAGTCGTCCTCCATCCCACTGACGTCCTTCGCCAGCT 3314  
DB 1172 TCCTGAGGATCAGCCTCACCCACCCAGTCGTCCTCCATCCCACTGACGTCCTTCGCCAGCT 1231  
QY 3315 TGGTAGCGGTGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGCCCGCAGACCTGTCT 3374  
DB 1232 TGGTAGCGGTGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGCCCGCAGACCTGTCT 1291  
QY 3375 ATTATTTCATTTGGGACAAGACAGACGCTCTACAACAGAGGTTTGGGCTTTTCAGAG 3434  
DB 1292 ATTATTTCATTTGGGACAAGACAGACGCTCTACAACAGAGGTTTGGGCTTTTCAGAG 1351  
QY 3435 CCTTTGTTCCGTGGGTTATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAA 3494  
DB 1352 CCTTTGTTCCGTGGGTTATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAA 1411  
QY 3495 GAACAACAGTCTGCAGGGCTATGAATTCAGCTTCAAGCTTGGAGGATGGAGCCTAG 3554  
DB 1412 GAACAACAGTCTGCAGGGCTATGAATTCAGCTTCAAGCTTGGAGGATGGAGCCTAG 1471  
QY 3555 ACAAACATCATCCCTCAACATTCAGAGTGGTATCCTGTCACAAAGGGAATGGGAGAAC 3614  
DB 1472 ACAAACATCATCCCTCAACATTCAGAGTGGTATCCTGTCACAAAGGGAATGGGAGAAC 1531  
QY 3615 AGTTTGTCTCAGCAGCCTCCTGTTCATTTGGGAGCATCATGGCAATGGCGCCGGAGAA 3674  
DB 1532 AGTTTGTCTCAGCAGCCTCCTGTTCATTTGGGAGCATCATGGCAATGGCGCCGGAGAA 1591  
QY 3675 GCATCTCCTGCCCCAGCTGCAACGGCTTGTGACGGCAACAGCTCCTGGCCCCCAGTGG 3734  
DB 1592 GCATCTCCTGCCCCAGCTGCAACGGCTTGTGACGGCAACAGCTCCTGGCCCCCAGTGG 1651  
QY 3735 CCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAGGA 3794  
DB 1652 CCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAGGA 1711  
QY 3795 TCTTCCCTCTGGAAATGTCAACAATCCTTAGAGCTGAGGAATAAGATTTCAGACATA 3854  
DB 1712 TCTTCCCTCTGGAAATGTCAACAATCCTTAGAGCTG-----A 1750  
QY 3855 GTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGGGCCGTCTTCC 3914  
DB 1751 GTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGGGCCGTCTTCC 1810  
QY 3915 TTTCTGACAGCAACAGCCGGGCTTTTAAATCAAGTCCACTGTGGTGGTGAAGACC 3974  
DB 1811 TTTCTGACAGCAACAGCCGGGCTTTTAAATCAAGTCCACTGTGGTGGTGAAGACC 1870  
QY 3975 TTGTCAAGAACTCTGAGGTGGTGGGGGACAGGTGACCAGTGCCTCCCTTTGATGACA 4034

Db 1871 TTGTCAAGAACTCTGAGGTGGTTGGGGGACAGGTGACCAAGTGCCTCCCTTTGATGACA 1930  
QY 4035 CTCGCTGGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCCAGGGGTATTA 4094  
Db 1931 CTCGCTGGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCCAGGGGCATTA 1990  
QY 4095 CAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCAATGATCAGACGCATCGATC 4154  
Db 1991 CAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCAATGATCAGACGCATCGATC 2050  
QY 4155 AGAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCATATCAGCCCCGGCCACTCA 4214  
Db 2051 AGAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCATATCAGCCCCGGCCACTCA 2110  
QY 4215 GCTGTGATTCTGTCAATGATATTTCCAGGTAAGACTGGAGTGGCCCCACAGACTTAGCCA 4274  
Db 2111 GCTGTGATTCTGTCAATGATATTTCCAGGTTACCTGGAGTGGCCCCACAGACTTAGCCA 2170  
QY 4275 TCAACCCCAATGGACAACTCACTTTATGTCCTCGAACAAATGAGTGCCTGCAAAATCTCTG 4334  
Db 2171 TCAACCCCAATGGACAACTCACTTTATGTCCTCGAACAAATGAGTGCCTGCAAAATCTCTG 2230  
QY 4335 AAAACCCAGGTGCGCATTTGTCGGGGAGGCCCATGCACTGCCAGGTCCCTGGCAATTG 4394  
Db 2231 AAAACCCAGGTGCGCATTTGTCGGGGAGGCCCATGCACTGCCAGGTCCCTGGCAATTG 2290  
QY 4395 ACCACTTCCTGCTAAGCAAGGTGGCCATCCACGCAACCTCGAGTCAGCCACCGCTTTGG 4454  
Db 2291 ACCACTTCCTGCTAAGCAAGGTGGCCATCCACGCAACCTCGAGTCAGCCACCGCTTTGG 2350  
QY 4455 CTGTTTCACACAATGGGGTCCCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCCGA 4514  
Db 2351 CTGTTTCACACAATGGGGTCCCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCCGA 2410  
QY 4515 TCAGGCAGGTCAACCACTAGTGGAGAGATCTACTCGTTGCTGGGGCCCCCAGTGGCTGTG 4574  
Db 2411 TCAGGCAGGTCAACCACTAGTGGAGAGATCTACTCGTTGCTGGGGCCCCCAGTGGCTGTG 2470  
QY 4575 ACTGTAAAAATGATGCCAACTGTGATTTCTGGAGACGATGGTTATGCCAAGGATG 4634  
Db 2471 ACTGTAAAAATGATGCCAACTGTGATTTCTGGAGACGATGGTTATGCCAAGGATG 2530  
QY 4635 CAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGAGCTCTACGTGGCCG 4694  
Db 2531 CAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGAGCTCTACGTGGCCG 2590  
QY 4695 ACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAAACAAGCCTTTCCTCAACACCCAGA 4754  
Db 2591 ACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAAACAAGCCTTTCCTCAACACCCAGA 2650  
QY 4755 ACATGTATGAGCTGTCTTCAACCAATTGACCAAGAGCTCTATCTGTTTGATACCCCGGCA 4814  
Db 2651 ACATGTATGAGCTGTCTTCAACCAATTGACCAAGAGCTCTATCTGTTTGATACCCCGGCA 2710  
QY 4815 AGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACTTACACTG 4874  
Db 2711 AGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACTTACACTG 2770  
QY 4875 GGGACGGCGACATCACTCATCAAGACAACAAATGGCAACATGGTAAATGTCCGCCGAG 4934  
Db 2771 GGGACGGCGACATCACTCATCAAGACAACAAATGGCAACATGGTAAATGTCCGCCGAG 2830  
QY 4935 ACTCTACTGGGATGCCCTCTGGCTGGTCCAGATGGCCAGGTGTACTGGGTGACCA 4994  
Db 2831 ACTCTACTGGGATGCCCTCTGGCTGGTCCAGATGGCCAGGTGTACTGGGTGACCA 2890  
QY 4995 TGGGCACCAACAGTGCCTCAAGAGTGTGACCAACAAGGACACGAGTTGGCCATGATGA 5054  
Db 2891 TGGGCACCAACAGTGCCTCAAGAGTGTGACCAACAAGGACACGAGTTGGCCATGATGA 2950  
QY 5055 CATAACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACAACAT 5114

Db 2951 CATAACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACAACAT 3010  
QY 5115 TTTATGAGTACGACAGACTTTGGCCGCTTGACAAATGTGACCTTCCCTACTGGCCAGGTGA 5174  
Db 3011 TTTATGAGTACGACAGACTTTGGCCGCTTGACAAATGTGACCTTCCCTACTGGCCAGGTGA 3070  
QY 5175 GCAGTTTCCGAAGTATACAGACAGTTTCAAGTGCATGTCCAGGTAGACACTCCAGCAAGG 5234  
Db 3071 GCAGTTTCCGAAGTATACAGACAGTTTCAAGTGCATGTCCAGGTAGACACTCCAGCAAGG 3130  
QY 5235 ATGATGTACCAATAACCAACCAACCTGTCTGCTCAGGCGCTTCTACACACTGCTGGCCA 5294  
Db 3131 ATGATGTACCAATAACCAACCAACCTGTCTGCTCAGGCGCTTCTACACACTGCTGGCAAG 3190  
QY 5295 ACCAAGTCCGGAACAGCTACTACATCGGGGCGGATGGCTCCTTGGGCTGCTGGCCA 5354  
Db 3191 ACCAAGTCCGGAACAGCTACTACATCGGGGCGGATGGCTCCTTGGGCTGCTGGCCA 3250  
QY 5355 ACGGCATGGAGTGGCGCTGCAGACTGAGCCCCACTTGTGGTGGCACCCGTCAACCCCCA 5414  
Db 3251 ACGGCATGGAGTGGCGCTGCAGACTGAGCCCCACTTGTGGTGGCACCCGTCAACCCCCA 3310  
QY 5415 CCGTGGGCAAGAGGAATGTCAAGTGCCTGCCATCGAACACGGCCTCAACCTGGTGGAGTGGC 5474  
Db 3311 CCGTGGGCAAGAGGAATGTCAAGTGCCTGCCATCGAACACGGCCTCAACCTGGTGGAGTGGC 3370  
QY 5475 GCCAGCGCAAAGAGCAGGCTCGGGGCCAGGTCACCTGTTTGGGCGCGGCTCGGGTGC 5534  
Db 3371 GCCAGCGCAAAGAGCAGGCTCGGGGCCAGGTCACCTGTTTGGGCGCGGCTCGGGTTC 3430  
QY 5535 ACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATG 5594  
Db 3431 ACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATG 3490  
QY 5595 ATGACCAACCGCAAGTTCAACCTTTCGGATTCTGTAGACAGGCGGGCGGCCAGCCTCT 5654  
Db 3491 ATGACCAACCGCAAGTTCAACCTTTCGGATTCTGTAGACAGGCGGGCGGCCAGCCTCT 3550  
QY 5655 GGTCAACCCAGCAGCAGGCTGAATGGTGTCAACCTGACATCTCCCTGGGGTTACATTG 5714  
Db 3551 GGTCAACCCAGCAGCAGGCTGAATGGTGTCAACCTGACATCTCCCTGGGGTTACATTG 3610  
QY 5715 CTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACAGGCGGGCGGCATCA 5774  
Db 3611 CTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACAGGCGGGCGGCATCA 3670  
QY 5775 CATCCAGGATCTTCGCTGATGGGAAGACATGGAGTACACATCTTAGAGAAGCAGGTT 5828  
Db 3671 CATCCAGGATCTTCGCTGATGGGAAGACATGGAGTACACATCTTAGAGAAGCAGGTT 3730  
QY 5829 CCATGGTGTCTACTACACAGCCAGAGGCGAGTATCTTTGAGTTTCGACAAAGATGACC 5888  
Db 3731 CCATGGTGTCTACTACACAGCCAGAGGCGAGTATCTTTGAGTTTCGACAAAGATGACC 3790  
QY 5889 GCCTCTCTTCTGTGACGATGCCCAACGTTGGCGGCGAGACACTAGAGACCATCCGCTCAG 5948  
Db 3791 GCCTCTCTTCTGTGACGATGCCCAACGTTGGCGGCGAGACACTAGAGACCATCCGCTCAG 3850  
QY 5949 TGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACT 6008  
Db 3851 TGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACT 3910  
QY 6009 TCACTGAGGATGGGCACCTCTTTCACACCTTCTACCTGGGCACTGGCCCGCAGGTTGATAT 6068  
Db 3911 TCACTGAGGATGGGCACCTCTTTCACACCTTCTACCTGGGCACTGGCCCGCAGGTTGATAT 3970  
QY 6069 ACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCACCAAGGTCAGTT 6128  
Db 3971 ACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCACCAAGGTCAGTT 4030  
QY 6129 TCACCTATGACGAGACGGCAGGCTGCTGAAGACCAATCAACCTACAGAAATGAGGGCTTCA 6188  
Db 4031 TCACCTATGACGAGACGGCAGGCTGCTGAAGACCAATCAACCTACAGAAATGAGGGCTTCA 4090







QY 6047 GGCACCTGGCCGAGGGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTC 6106  
Ddb 3673 GGCACCTGGCCGAGGGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTC 3732  
QY 6107 TATGACACCAACCAAGGTCAGTTTCACCTATGACGAGACGGCAGGCTGCTGAAGACCATC 6166  
Ddb 3733 TATGACACCAACCAAGGTCAGTTTCACCTATGACGAGACGGCAGGCTGCTGAAGACCATC 3792  
QY 6167 AACCTACAGAATGAGGGCTTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCTGATT 6226  
Ddb 3793 AACCTACAGAATGAGGGCTTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCTGATT 3852  
QY 6227 GACCGACAGATCTTCCGCTTCACCTGAGGAAGGCATGGTCAACGCCCTTTTGACTACAAC 6286  
Ddb 3853 GACCGACAGATCTTCCGCTTCACCGAGGAAGGCATGGTCAACGCCCTTTTGACTACAAC 3912  
QY 6287 TATGACAAACAGCTTCCGGGTGACCAAGCATGTCAGGCTGTGATCAACGAGACCCCACTGCC 6346  
Ddb 3913 TATGACAAACAGCTTCCGGGTGACCAAGCATGTCAGGCTGTGATCAACGAGACCCCACTGCC 3972  
QY 6347 ATTGATCTCTATCGCTATGATGATGTGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGT 6406  
Ddb 3973 ATTGATCTCTATCGCTATGATGATGTGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGT 4032  
QY 6407 GTCATTTACTATGACATTAACCAAGATCATCAACACAGCTGTATGATGAGATCTTCCGCTCGCTCATGTAC 6466  
Ddb 4033 GTCATTTACTATGACATTAACCAAGATCATCAACACAGCTGTATGATGAGATCTTCCGCTCGCTCATGTAC 4092  
QY 6467 TTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAC 6526  
Ddb 4093 TTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAC 4152  
QY 6527 TGGATGACCGTCCAGTATGATAAACATGGGGCGAGTAGTGAAGAGGAGTGAAGGTAGGA 6586  
Ddb 4153 TGGATGACTGTCCAGTATGATAAACATGGGGCGAGTAGTGAAGAGGAGTGAAGGTAGGA 4212  
QY 6587 CCCTACGCCAATACCACTCGCTACTCCTATGATGATGATGCTGACGGCCAGCTGCAGACA 6646  
Ddb 4213 CCCTACGCCAATACCACTCGCTACTCCTATGATGATGATGCTGACGGCCAGCTGCAGACA 4272  
QY 6647 GTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCAC 6706  
Ddb 4273 GTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCAC 4332  
QY 6707 TTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGC 6766  
Ddb 4333 TTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGC 4392  
QY 6767 ATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGCTTCTGAGGCAGCGG 6826  
Ddb 4393 ATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGCTTCTGAGGCAGCGG 4452  
QY 6827 GCGGTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTACAACCGGGCT 6886  
Ddb 4453 GCGGTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTACAACCGGGCT 4512  
QY 6887 GGCAGCTGGAGTGCAGGTACCGCTACGATGGCTGGGGCGGCGGTGTCCAGCAAGAGC 6946  
Ddb 4513 GGCAGCTGGAGTGCAGGTACCGCTACGATNNNTNNNTNNNTNNNTNNNTNNNTNNNTNNNT 4572  
QY 6947 AGCCACAGCCACCACTGCAGTTCTTCTATGCAGACCTGACCAACCCCAAGGTCAAC 7006  
Ddb 4573 AGCCACAGCCACNN 4632  
QY 7007 CACCTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGAAGGACAC 7066  
Ddb 4633 NNN 4692  
QY 7067 CTCTTTGCCATGGAGCTGAGCAGTGGTGTAGTGGTTTACATAGCTTGTGACAACATCGGG 7126  
Ddb 4693 CTCTTTGCCATGGAGCTGAGCAGTGGTGTAGCAGTTTACATAGCTTGTGACAACATCGGG 4752  
QY 7127 ACCCTCTTGCTGTCTTAGTGGAAACAGGTTTGATGATCAAGCAAACTCTGTACACAGCC 7186

Ddb 4753 ACCCTCTTGCTGTCTTTAGTGGAAACAGGTTTGATGATCAAGCAAAATCTGTACACAGCC 4812  
QY 7187 TATGGGAGATCTACATGATACCAACCCCAACTTTAGATCATATAGGCTACCATGGT 7246  
Ddb 4813 TATGGGAGATCTACATGATACCAACCCCAACTTTAGATCATATAGGCTACCATGGT 4872  
QY 7247 GGCCTCTATGATCCACTACCAAGCTTGTCCACATGGGCGCGAGATTATGATGTGCTG 7306  
Ddb 4873 GGCCTCTATGATCCACTACCAAGCTTGTCCACATGGGCGCGAGATTATGATGTGCTG 4932  
QY 7307 GCCGACGCTGGACTAGCCAGACCAAGCTGTGGAAGCACCTTAGTAGCAGCAACGTC 7366  
Ddb 4933 GCCGACGCTGGACTAGCCAGACCAAGCTGTGGAAGCACCTTAGTAGCAGCAACGTC 4992  
QY 7367 ATGCCTTTTAATCTCTATATGTTCAAAAACAACCAACCCCATCAGCAACTCCCAGGACATC 7426  
Ddb 4993 ATGCCTTTTAATCTCTATATGTTCAAAAACAACCAACCCCATCAGCAACTCCCAGGACATC 5052  
QY 7427 AAGTGCTTCATGACAGATGTTAAACAGCTGGCTGCTCACTTTGGATTCCAGCTACACAAC 7486  
Ddb 5053 AAGTGCTTCATGACAGATGTTAAACAGCTGGCTGCTCACTTTGGATTCCAGCTACACAAC 5112  
QY 7487 GTGATCCCTGTTTATCCCAACCAAGACATGGATGCCATGGAAACCTCCTACGAGCTCATC 7546  
Ddb 5113 GTGATCCCTGTTTATCCCAACCAAGACATGGATGCCATGGAAACCTCCTACGAGCTCATC 5172  
QY 7547 CACACACAGATGAAAAACGAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGT 7606  
Ddb 5173 CACACACAGATGAAAAACGAGGAGTGGGACAAACAGCAAGTCTATCCTTGGGGTACAGTGT 5232  
QY 7607 GAAGTACAGAAGCAGCTCAAGGCCCTTTGTCACTTAGAACGGTTTGACCAAGCTCTATGGC 7666  
Ddb 5233 GAAGTACAGAAGCAGCTCAAGGCCCTTTGTCACTTAGAACGGTTTGACCAAGCTCTATGGC 5292  
QY 7667 TCACAATCACCAGCTGCCAGAGGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCA 7726  
Ddb 5293 TCACAATCACCAGCTGCCAGAGGCTCCAAAGACCAAGAGTTTGCATCCAGCGCTCA 5352  
QY 7727 GTCTTTGGCAAGGGGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAAGACATCATC 7786  
Ddb 5353 GTCTTTGGCAAGGGGTCAAGTTTGCTTGAAGGATGGCCGAGTGACCAAGACATCATC 5412  
QY 7787 AGTGTGGCCAAATGAGGATGGCGAAGGGTTGTGCCATCTTGAACCATGCCCCACTACCTA 7846  
Ddb 5413 AGTGTGGCCAAATGAGGATGGCGAAGGGTTGTGCCATCTTGAACCATGCCCCACTANCNA 5472  
QY 7847 GAGAACCTGCACCTTCAACATGATGGGTGATACCACTTACTTTGTGAAACCAAGACCT 7906  
Ddb 5473 GAGAACCTGCACCTTCAACATGATGGGTGATACCACTTACTTTGTGAAACCAAGACCT 5532  
QY 7907 TCAGAAAGGTGACCTGGCCATCTGGCCCTCAGTGGGGGGCGGCAACCTGGAGAATGGG 7966  
Ddb 5533 TCAGAAAGGTGACCTGGCCATCTGGCCCTCAGTGGGGGGCGGCAACCTGGAGAATGGG 5592  
QY 7967 GTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACA 8026  
Ddb 5593 GTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACA 5652  
QY 8027 GACATCCAGCTCCAGTACGGGGCCTGTGCTTGAACACACAGCTACGGGACACGTTGGAT 8086  
Ddb 5653 GACATCCAGCTCCAGTACGGGGCCTGTGCTTGAACACACAGCTACGGGACACGTTGGAT 5712  
QY 8087 GAGGAGAAAGGACGGGTCTGGAGCTGGCCCGGAGAGAGCCGTGCGCCAAGCTGGGCC 8146  
Ddb 5713 GAGGAGAAAGGACGGGTCTGGAGCTGGCCCGGAGAGAGCCGTGCGCCAAGCTGGGCC 5772  
QY 8147 CGCAGCAGCAGAGACTCGCGGAAGGGAGGAAAGCCCTGCGGCCCTGGACAGAGGGGGAG 8206  
Ddb 5773 CGCAGCAGCAGAGACTCGCGGAANNNGAGGAAGCCCTGCGGCCCTGGACAGAGGGGGAG 5832  
QY 8207 AAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAGGCTACGACGGCTTTTTCGTGATCTCT 8266

Db 5833 AAGCAGCAGGTCTGTAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCT 5892

QY 8267 GTCGAGCAGTACCAGAACTGTCTCAGACAGCGGCCAAACAACATCCACTTTCATGAGACAGGC 8326

Db 5893 GTCGAGCAGTACCAGAACTGTCTCAGACAGCGGCCAAACAACATCCACTTTCATGAGACAGGC 5952

QY 8327 GAGATGGGCGGAGGTGA 8344

Db 5953 GAGATGGGCGGAGGTGA 5970

RESULT 3

LOCUS AY405420 5094 bp DNA linear GSS 12-DEC-2003

DEFINITION Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY405420

VERSION AY405420.1 GI:39761394

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5094)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 5094)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source

1..5094

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

<1..>5094

/locus\_tag="HCM2218"

gene

ORIGIN

Query Match 28.2%; Score 2353; DB 29; Length 5094;

Best Local Similarity 67.0%; Pred. No. 0;

Matches 3442; Conservative 0; Mismatches 1647; Indels 45; Gaps 6;

QY 2912 TTTGACTTGGTGACAAATGGCGGCATCTCCATCATCTCGGGTTCGAGCGGCACCTTTC 2971

Db 2 TTTGACTTGGTGACAAATGGTGGGGCTCTCTAACTTTGGTATTGAAACGATCCCCATTC 61

QY 2972 ATCACACAGGACACACCCTGTGGTGGCCATGGGATCGCTTCTTGTCTATGGAAACCATC 3031

Db 62 CTCACCTAGTATCATACTGTGTGGATTCATGGAAATGCTTTTATGTGATGGATACCCCTA 121

QY 3032 ATCATGAGACATGAGGAGAAATGAGATTCACAGCTGTGACCTGAGCAATTTTGGCCGCCCC 3091

Db 122 GTCATGAAGAAAGAGAGAAATGACATTCACAGCTGTGATCTGAGTGGATTCGTGAGGCCA 181

QY 3092 AACCCAGTCTCTCTCCATCCCCACTGACGTCTTTCGCCAGTCTCTGTGACAGAAAGGC 3151

Db 182 AATCCCATCATTTGTGTATCATCACCTTTATCCACCTTTTTCAGATCTTCTCCTGAAGACAGT 241

QY 3152 CCCATTGTCCGGAAATTCAGGCTTTTCAGGAGGAAATCTCTATCTCTGGCTGCAAGATG 3211

Db 242 CCCATCATTTCCGAAACACANGTACTCCACGAGGAAACTACAATTCAGGAACAGATTG 301

QY 3212 AGGCTGAGCTACCTGAGCAGCCCGGACCCCTGGCTACAAAATCTGTCTCTGAGGATCAGCCTC 3271

Db 302 AAACCTCTCTACTTGAAGTCCAGAGCTGCAGGGTATAAGTCAGTTCTCAAGATCACCATG 361

QY 3272 ACCCACCCGACCATCCCCCTTCAACCTCATGAAGGTGCACCTCATGGTAGCGGTGGAGGC 3331

Db 362 ACCCAGTCTATTATTCATTTAATTAATGAAGTTTCATCTTATGGTAGCTGTAGTAGGA 421

QY 3332 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCCAGACCTGTCTCTATTATTTTCATTTGGAC 3391

Db 422 AGACTCTTCCAAAAGTGGTTTCTGCTCACCAAACCTTGGCTATATCTTTCATATGGGAT 481

QY 3392 AAGACAGACGTCTACAACCAAGAGGTGTTTGGGCTTTTCAGAAGCCTTTGTTCCGTGGGT 3451

Db 482 AAAACAGATGCATATAATCAGAAAAGTCTATGCTCTATCTGAAGCTGTTGTGTCAAGTTGA 541

QY 3452 TATGAATATGAATCCTGCCCCAGATCTAATCCTGTGGGAAAAGAAACAACAGTGTGCAG 3511

Db 542 TATGAGTATCAGTCTGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 601

QY 3512 GGCTATGAATTGACGCGTCCAAAGCTTGAGGATGGAGCCCTAGACAAAACATCATGCCCTC 3571

Db 602 GGCTATGAATTGAGTCTGCTCAACATGGTGGTGGCTGGACATTAGATAAAACATCACGTGCTG 661

QY 3572 AACATTCAAAGTGGTATCCTGCAAAAAGGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 3631

Db 662 GATGTACAGAACGGTATACCTGTACAAGGGAACGGGGAAAACCAAGTTTCATCTCCCAGCAG 721

QY 3632 CCTCCTGTCTATGGGAGCATCATGGCAATGGCGCCGCGGAGAGCATCTCCTGCCCCCAGC 3691

Db 722 CCTCCAGTCTGAGTAGCATCATGGCAATGGCGAAGGCGCAGCATTTCTGCCCCCAGT 781

QY 3692 TGCAACGGCCTTGTGACGGCAACAAGCTCTCTGGCCCCCAGTGGCCCTCACCTGTGGCTCT 3751

Db 782 TGCAATGGTCAAGCTGATGGTAACAAGTTACTGGCCCCCAGTGGCGCTAGCTTGTGGGATC 841

QY 3752 GACGGGAGCCTCTATGTGGTGTATTCAACTACATTAGAGGATCTTCCCCTCTGGAAT 3811

Db 842 GATGGCAGTCTGTACGTAGGCGATTCAACTATGTGCGGGGATATTCCTTCTTGGAAT 901

QY 3812 GTCACCAACATCCTAGAGCTGAGGAATAAGATTTTCAGACATAGTCCAGCAGCAC 3871

Db 902 GTAACAAGTCTTAGAACTAAGCA-----GCAACCCAGCTCAT 940

QY 3872 AATACTACTGTGGCCACAGACCCCATGAGTGGGGCCGCTCTTCTTCTGACAGCAACAGC 3931

Db 941 AGATACTACTGTGCAACGGATCCAGTCCAGGAGATCTGTACGTTTCTGACACAACACC 1000

QY 3932 CGGCGGTCTTTAAAAATCAAGTCCACTGTGGTGGTGAAGACCTTGTCAAGAACTCTGAG 3991

Db 1001 CGCAGAAATTTATCGCCCCAAAGTCACTACGGGGGCAAAAGACTTGACTAAAAATGCAGAA 1060

QY 3992 GTGGTTGCGGGACAGGTGACCAAGTCCCTTTCCTTGTATGACACTCGCTGCGGGGATGGT 4051

Db 1061 GTCGTGCGAGGACAGGGGAGCAATGCCCTTCGTTTGACGAGCGAGATGTGGGATGGA 1120

QY 4052 GGAAGGCCACAGAAAGCCACACTCAACAATCCAGGGGTATTACAGTGGACAAGTTTGGG 4111

Db 1121 GGAAGGCCGTGGAAGCCACACTCATGAGTCCCAAGGAATGGCAGTTGTAAAGATGGA 1180

QY 4112 CTGATCTACTTCGTGGATGGCACCATGATCAGACGATCAGATCGAATGGGATCATCTCC 4171

Db 1181 TTAATCTACTTTGTTGATGGAACCATGATTAGAAAGTTGACCAAAATGGAATCATATCA 1240

QY 4172 ACCCTGCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGTGATTCTGTGATG 4231

Db 1241 ACTCTTCTGGCTCTAACGATTTGACTTCAGCCAGACCTTTAACTTGTGACACCAGCATG 1300

QY 4232 GATATTTCCAGGTAAGACTGGAGTGGCCCAAGACTTAGCATCAACCCCAATGGACAAC 4291

Db 1301 CATATCAGCCAGGTACGTCTGGAATGGCCCACTGACCTAGCATTAACCCCTATGGATAAC 1360



Db	3518	ATGACCTATACGAAGCACTTTGATGTCTCATGGCCGTATCAAGGAGATTCAATATGAGATA	3577
QY	6509	TTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAG	6568
Db	3578	TTCAGGTCGCTCATGTACTGGATTACAATTCAGTATGATAACATGGGTCCGGTAAACCAAG	3637
QY	6569	AAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGAGTATGATGCT	6628
Db	3638	AGAGAGATTAAATAGGCCCTTTTGCCAACACCACCAAAATATGCTTATGAATATGATGTT	3697
QY	6629	GACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTTGCGCTACAGCTACGAC	6688
Db	3698	GATGGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAAATGTGGCGGTACAACATACGAT	3757
QY	6689	CTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGG	6748
Db	3758	CTGAATGGAACCTCCATTTACTGAACCCCAAGTAAACAGTGCAGCTCTGACACCCCTTCGC	3817
QY	6749	TATGACATCCGCGACCGCATCACTCGGCTGGGTGAGCTGCAATACAAGATGGATGAGGAT	6808
Db	3818	TATGACCTGCGAGACAGAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGAT	3877
QY	6809	GGCTTCTTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATC	6868
Db	3878	GGTTTCTTACGTCAAAGGGCACGGAATCTTTGATATATAGCTCCAAGGGCTTTAACT	3937
QY	6869	AAGGCCTACAACCCGGGTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCTGGGCGG	6928
Db	3938	CGAGTTTACAGTAAAGCAGTGGCTGGACAGTGTCTACCGTTATGACGGCCTGGGAAG	3997
QY	6929	CGCGTGTCCAGCAAGCAGCACCCACAGCCACCCTCCAGTTCTTCTATGCAGACCTGACC	6988
Db	3998	CGTGTCTTAGCAAAACCAAGTCTAGGACAGCACCTGCAGTTTTTTTATGCTGACTTAACT	4057
QY	6989	AACCCCAACCAAGGTCAACCACTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTAC	7048
Db	4058	TATCCCACTAGGATTACTCATGTCTACAACCAATTCGAGTTCAGAAATTACCTCCCTGAT	4117
QY	7049	TACGACTTGCAAGGACACCTCTTTTGCCATGGAGCTGAGCAGTGGTGATGAGTTTACATA	7108
Db	4118	TATGATCTCCAAGGACATCTTTTTTGCCATGGAATCAGCAGTGGGATGAATTCATATT	4177
QY	7109	GCTTGTGACAAACATCGGACCCCTCTTTGCTGTCTTTAGTGGAAACAGTTTGATCATCAAG	7168
Db	4178	GCATCGGATAACACAGGGACACCACTGGCTGTGTTCAGTAGCAATGGGCTTATGCTGAAA	4237
QY	7169	CAAAATCCTGTACACAGCTATGGGGAGATCTACATGGATACCAACCCCACTTTCAGATC	7228
Db	4238	CAGATTTCAGTACACTGCATATGGGAAATCTATTTGACTCTAATATTGACTTTCAACTG	4297
QY	7229	ATCATAGGCTACCATGTRGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGG	7288
Db	4298	GTAATTGGATTTCATGTGGCCTGTATGACCCACTCACAAAATTAATCCACTTTGGAGAA	4357
QY	7289	CGAGATTATGATGTGTGGCCGGACGCTGGACTAGCCAGACCAACGAGCTGTGGAAGCAC	7348
Db	4358	AGAGATTATGACATTTTGGCAGGACGGTGGACAACACCTGACATAGAAATCTGGA--AA	4414
QY	7349	CTTAGTAGCAGCAACGTCATGCTCTTTTAATCTCTATATGTTCAAAAACAACACCCCATC	7408
Db	4415	AGAAATGGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACACCCCTGCA	4474
QY	7409	AGCAACTCCCAGGACATCAAGTGTCTTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTT	7468
Db	4475	AGCAAAATCCATGACGTTGAAGATTACATCACAGATGTTAACAGCTGGCTGGTGACATTT	4534
QY	7469	GGATTCCAGCTACACAACGTTGATCCCTGGTTATCCCAACCCAGACATGGATGCCATGGAA	7528
Db	4535	GGTTTCCATCTGCACAAATGCTATTCTCTGGATTCTCCCTGTTCCCAAATTTGATTTAACAGAA	4594
QY	7529	CCCTCCTACGAGCTCATCCACACACAGATGAAAACCGCAGGAGTGGGACAACAGCAAGTCT	7588

Db	4595	CCTTCTTACGA-----ACTTGTGAAGAGTCAGCAGTGGGATGATATACCGCCC	4642
QY	7589	ATCCTCGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGTCACTTAGAACGG	7648
Db	4643	ATCTTCGGAGTCAGCAGCAAGTGGCGGCAGGCCAAGGCCCTTCCTGTGCTGGGGGAAG	4702
QY	7649	TTTGACCAGCTCTATGGCTCCACAATCACCAAGTGCAGCAGGCTCCAAAGACCAAGAAG	7708
Db	4703	ATGGCCGAGGTGAGGTGAGCCGGCGCCGGGCCGGC--GGCGCGCAGTCTCTGCTGTGG	4759
QY	7709	TTTGCAATCCAGCGGCTCAGTCTTTTGGCAAGGGGCTCAAGTTTGCTTGAAGGATCGGCCGA	7768
Db	4760	TTTCGCCACGGTCAAGTCGCTGATCGGCAAGGGCGTCATGCTGGCCGTCAGCCAGGCCCGC	4819
QY	7769	GTGACCACAGACATCATCAGTGTGGCCAATGAGGATGGCGAAGGTTGCTGCGATCTTTG	7828
Db	4820	GTGCAGACCAACGTGCTCAACATCGCCAACGAGGACTGCATCAAGGTGGCGCCGTGCTC	4879
QY	7829	AACCATGCCACTACCTAGAGAACCTGCACCTTACCATTTGATGGGTGGATACCCATTAC	7888
Db	4880	AACAACGCCCTTCTACCTGGAGAACCTTGCACCTTCAACCATCGAGGGCAAGGACACGACTAC	4939
QY	7889	TTTGTGAACACAGACCTTCAGAAAGGTGACCTTGCCCATCCTGGGCCCTCAGTGGGGGCGG	7948
Db	4940	TTTCATCAAGACCAACACGCCCGAGAGCGACCTTGGSCACGCTGCGGTTGACCCAGCGCCGC	4999
QY	7949	CGAACCCCTGGAGATGGGGTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAATGGC	8008
Db	5000	AAGGCGCTGGAGAACGGCATCAACGTCACGGTGTGCGAGTCCACCACGGTGTGTAACGGC	5059
QY	8009	AGGACTAGACGCTACACAGACATCCAGCTCCAGT	8042
Db	5060	AGGACCGCGCAGGTTCCGCGACGCTGGAGATGCAGT	5093

RESULT 4	
AY413477	
LOCUS	AY413477 3190 bp DNA linear GSS 12-DEC-2003
DEFINITION	Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION	AY413477
VERSION	AY413477.1 GI:39769439
KEYWORDS	GSS.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3190)
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 3190)
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1. .3190 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>3190 /locus tag="HCM4903"
gene	

ORIGIN

Query Match 28.1%; Score 2347.8; DB 29; Length 3190;									
Best Local Similarity 83.3%; Pred. No. 0;									
Matches 2612; Conservative 0; Mismatches 474; Indels 48; Gaps 2;									
QY	2161	CACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACAGACTGTTCTATCGAGATCTG	2220						
Db	105	CATGGGCACCTGCATCTGCAACCCCGGCTACAAGGCGAGAGCTGCGAGGAAGATCTG	164						
QY	2221	TGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGCGACCTGCCCGCTGCCAGGATGG	2280						
Db	165	TGCCGCTGACTGTGGTGGCCATGGCGTCTGCGTAGAGGCACCTGCCGCTGTGAGGATGG	224						
QY	2281	CTGGATGGGGCGACGCTGCGACACGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGAC	2340						
Db	225	TTGGATGGGGCCGCGATGCGACCAACGGGCTGCCACCCACGCTGTGCAGAACACGGGAC	284						
QY	2341	CTGCCGCGACGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCAACACTGCACCATCGC	2400						
Db	285	CTGCCGGACGGCAAGTGCGAATGCAGCCCGGCTGGAATGGAGAGCACTGCACCATC--	342						
QY	2401	TCACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAG	2460						
Db	343	-----GNN	377						
QY	2461	ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGAGCTGG	2520						
Db	378	NN	437						
QY	2521	CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAGACAATGATGGAGATGGCCT	2580						
Db	438	NN	497						
QY	2581	GGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG	2640						
Db	498	GGTGGACTGCATGGACCCCTGACTGCTGCCCTCAGCCCTCTCTGTCAATGTCACACCCGCTGTG	557						
QY	2641	CCTTGGCTCCCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGCA	2700						
Db	558	CCTAGGCTCCCCCTGACCCCTGGACATCATCAAGAGACACAGGCCCTGTATCCCAGCA	617						
QY	2701	GAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGCGAGGACAGCACGCACAT	2760						
Db	618	GAACCTGAACCTCCTTCTACGATCGAATCAAGTTCCTCGTGGGCGAGGACAGCACACAG	677						
QY	2761	AATCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCTGTCGGCCAAAGTGAT	2820						
Db	678	CATCCCGGGGAGAACCCCTTTGACGGAGGGCATGCTTGTGTCTATCCGTGGACAAAGTGAT	737						
QY	2821	GACATCAGATGGAAACCCCTTGTGTTGGTGTGAACATCAGTTTTGTCAATAACCCCTCTCTT	2880						
Db	738	GACGTGAGATGGGACCCCATTTGTTGGCGTGAACATCAGTTTTTCATCAATAACCCCTCTCTT	797						
QY	2881	TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC	2940						
Db	798	TGGATACACAATCAGCAGGCAAGATGGCAGCTTTGACCTGGTCACAAAATGGCGGCATCTC	857						
QY	2941	CATCATCCTGCGGTTGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCC	3000						
Db	858	CATCATCTTGAGATTCGAAAGGGCACCCCTTCATCACACAGGAGCATACCCCTCTGGCTACC	917						
QY	3001	ATGGGATCGCTTCTTGTGATGGAAACCATCATCATGAGACATGAGGAGAAATGAGATTCC	3060						
Db	918	CTGGGATCGCTTCTTGTGATGGAACCATCGTCTATGAGACACGAGGAGATGAGATCCC	977						
QY	3061	CAGCTGTGACCTGAGCAATTTTGCCCGCCGCCAACCCAGTCTCTCTCCATCCCCACTGAC	3120						
Db	978	CAGCTGTGACCTGAGCAACTTTGCCCCGTCCCAACCCCTGTGGTCTCTCTCCATCCCCACTGAC	1037						
QY	3121	GTCCCTTCGCCAGCTCCTGTGCAGAGAAAGCCCCCATTTGTGCCGGAATTCAGGCTTTGCA	3180						
Db	1038	ATCCTTCGCCAGCTCCTGTGTGCTGAGAAAGGCCCATTTGTCCCAGAAATCCAGGCCCTGCA	1097						

QY	3181	GGAGGAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCC	3240
Db	1098	AGAGGAATCGTCATCGCTGGCTGCAAGATGAGGTTGAGCTACCTGAGCAGCCGCACTCC	1157
QY	3241	TGGCTACAAATCTGTCTCTGAGGATCAGCCTCACCCACCCGACCATCCCCTTCAACCTCAT	3300
Db	1158	TGGCTATAAGTCTGTCTCTGAGGATCAGTCTCACACACCCACCATCCCCTTCAACCTCAT	1217
QY	3301	GAAGTGCACCTCATGGTAGCGGTGGAGGGCCGCTCTTCAGGAAGTGTTGCTGCAGC	3360
Db	1218	GAAGTGCACCTCATGGTGGCAGTCGAGGGTCGGCTCTTCAGGAAGTGTTTGCCTGCAGC	1277
QY	3361	CCCAGACCTGTCTTATTATTTCATTTGGGACAAGACAGACGTCTACAACCAAGAGGTGTT	3420
Db	1278	TCCCTGACCTGTCTTATTACCTCATCTGGGACAAGACAGATGTCTACAACCAAGAGGTGTT	1337
QY	3421	TGGGCTTTCAGAAGCCCTTGTTCCTGTGGGTTATGAATATGAATCCTGCCCAGATCTAAT	3480
Db	1338	TGGACTCTCAGAAGCCCTTGTTCCTGTGGGTTATAGTATGAATCCTGCCCAGATCTGAT	1397
QY	3481	CCTGTGGGAAAAAAGAACACAGTGTGCAGGGCTATGAAATTTGACGCTCCAAGCTTGG	3540
Db	1398	CCTGTGGGAGAAAAAGGACAGCGGTGCTGCAGGGCTATGAAATTTGATGCTCCAAGCTGG	1457
QY	3541	AGGATGGAGCCTTAGACAAAACATCATGCCCTCAACATTTCAAAAGTGGTATCCTGCACAAAGG	3600
Db	1458	GGGCTGGAGTCTGGATAAGCACCATGCCCTGAACATCCAGAGTGGCATCCTGCACAAAGG	1517
QY	3601	GAATGGGAGAACCCAGTTTGTGTCTCAGCAGCCTCCTGTCTATTTGGGAGCATCATGGSCAA	3660
Db	1518	GAATGGAGAGAACCCAGTTTGTGTCCAGCAGCCACCGGTCTATCGGAGCATCATGGSCAA	1577
QY	3661	TGGGCGCCGAGAGACATCTCCTGCCAGCTGCAACGGCCTTGCTGACGGCAACAAGCT	3720
Db	1578	TGGGCGCAGGAGAAAGTATCTCCTGCCCAGCTGCAATGGTCTTGTGATGSCAACAAAGCT	1637
QY	3721	CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA	3780
Db	1638	CCTGGCCCCCGTGGCCCTCACCTGGGCTCTGTATGGAGTCTCTACGTGGGAGACTTCAA	1697
QY	3781	CTACATTAGAAGGATCTTCCCCTCTCGAAATGTCAACCAACATCCTAGAGCTGAGGAATAA	3840
Db	1698	CTACATCCGCAGAAATCTTCCCCTCTGGAATGTCAACCAACATCCTGGAGATG-----	1749
QY	3841	AGATTTTCAGACATAGTCAAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAG	3900
Db	1750	-----AGTCACAGCCCAGCACACAAATACTACCTGGCTACAGACCCCATGAG	1796
QY	3901	TGGGCGCGTCTTCTTCTTGACAGCAACAGCCGCGGTCTTTAAATCAAGTCCACTGT	3960
Db	1797	TGGGCGCGTCTTCTTCTTGACACCAACAGCCGCGGTCTTCAAGGTCAAGTCCACCAC	1856
QY	3961	GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACAGTGCCT	4020
Db	1857	AGTGGTGAAGGACCTGGTGAAGAACTCCGAGGTGGTAGCAGGGACTGGTGACAGTGCCT	1916
QY	4021	CCCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGCCACAGAAAGCCACACTCACCAA	4080
Db	1917	CCCCTTTGATGATACCCGCTCGGAGATGGTGGGAAGCCACAGAAAGCCACGCTCACTAA	1976
QY	4081	TCCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTTCGTGGATGGCACCATGAT	4140
Db	1977	CCCAGGGGAATTACAGTGGACAAAGTTTGGGCTCATTTATTTTCGTGGACGSCACCATGAT	2036
QY	4141	CAGACGCATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATC	4200
Db	2037	CAGACGTGTTGATCAAAATGGAATCATCTCCACTTTGCTGGCTCCCAATGACCTCACCTC	2096
QY	4201	AGCCCGGCCACTCAGCTGTGATTCTGTCTCATGATATTTCCCAGGTAAGACTGGAGTGGCC	4260
Db	2097	GCCAGGCCCTCAGCTGTGACTCCGTCTCATGAGATTTCTCAGGTTCCGCTCGAGTGGCC	2156

QY 4261 CACAGACTTAGCCATCAACCCCAATGGACAACCTCACTTTATGTCTCCTCGACAACAATGTGGT 4320  
Db 2157 CACAGACTTAGCCATCAACCCCAATGGATAATTCTCTCATGTCTCCTCGACAACAATGTGGT 2216  
QY 4321 CCTGCAATCTCTGAAAACCCACAGGTGCGCATTTGTGCGGGAGGCCCATGCACTGCCA 4380  
Db 2217 CCTGCAATCTCCGAAAACCCACAGGTACGCATTTGTGCTGGAGGCCCATGCACTGCCA 2276  
QY 4381 GGTCCCTGGCATTGACCACTTCCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440  
Db 2277 GGTCCCTGGCATTGACCACTTCCTGCTAAGCAAGGTGGCCATCCATGCGACCCCTGGAGTC 2336  
QY 4441 AGCCACCGCTTTGGCTGTTTTCACAAATGGGTCCTGTATATTGCTGAGACTGATGAGAA 4500  
Db 2337 AGCCACTGCTTGCGCGTTTTCACAAATGGAGTCTTTGATACATCGCTGAGACCCGAGAGAA 2396  
QY 4501 AAAGATCAACCGCATCAGGCAGGTACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGC 4560  
Db 2397 AAAGATCAATCGCATCAGGCAGGTCACTACAAGTGGTGAGATCTCACTGGTTGCTGGTGC 2456  
QY 4561 CCCAGTGGGTGACTGTAATAATGATGCCAACTGTGATGTTTTTCTGGAGACGATGG 4620  
Db 2457 CCCAGTGGGTGACTGTAATAATGATGCCAACTGTGATGTTTTTCTGGAGATGATGG 2516  
QY 4621 TTATGCCAAGATGCAAAAGTTAAATACCCCATCTTCCCTGGCTGTGTGCTGATGGGA 4680  
Db 2517 TTACGCCAAGATGCAAAAGTTAAATACCCCATCTGCTTGGTGTGTGCTGATGGGA 2576  
QY 4681 GCTCTACGTGGCGGACCTTGGGAACATCCGAATTCGTTTATCCGGAAGAACAGCCTTT 4740  
Db 2577 GCTCTATGTGGCGGACCTTGGGAACATCCGAATTCGTTTATCCGGAAGAACAGCCTTT 2636  
QY 4741 CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAAGGAGCTCTATCTGTT 4800  
Db 2637 CCTGAACACTCAGAACATGTACGAGCTATCTCTCCCATCGACAGGAGCTGTACCTCTT 2696  
QY 4801 TGATACCAACCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAA 4860  
Db 2697 TGATACCAAGTGAAGCATCTGTACACTCAGAGCCTACCCACAGGGGACTACCTGTACAA 2756  
QY 4861 CTTACCTACACTGGGACGGGACATCACTACATCAAGAGTGTGACCAACAAGGACACGA 4920  
Db 2757 CTTCACTTACACAGGGACGGGACATCACTACATCAACGCAACAATGCAACATGGT 2816  
QY 4921 AATGTCCGCGAGACTCTACTGGGATGCCCCCTCTGCTGTGTGCTGCCAGATGGCCAGGT 4980  
Db 2817 GAACGTCCGCGAGACTCTACCGGATGCTCTCTGCTGTGTGCTGCCAGATGGCCAGGT 2876  
QY 4981 GTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACCAACAAGGACACGA 5040  
Db 2877 ATACTGGGTAAACCATGGGACCAACAGCGCACTCAGAAAGTGTGACCAACAAGGACACGA 2936  
QY 5041 GTTGGCCATGATGACATACCATGGGAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAA 5100  
Db 2937 GCTAGCCATGATGACCTACCATGGCAACTCTGGCCTCTTGGCAACCAAAAGCAATGAAA 2996  
QY 5101 CGGATGGACAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5160  
Db 2997 CGGGTGGACAACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3056  
QY 5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTGTATGATGATGATGATGATGATGATGATGATGAT 5220  
Db 3057 AACTGGCCAGGTGAGCAGTTTCCGAAGTGTATGATGATGATGATGATGATGATGATGATGAT 3116  
QY 5221 GACCTCCAGCAAGATGATGTACCAATACCAACCAACCTGTGTGCTCAGGCGCTTCTA 5280  
Db 3117 GACCTCAAGCAAGATGATGTACCAATACCAACCAACCTGTGTGCTCAGGCTGCTTCTA 3176  
QY 5281 CACACTGCTGCAAG 5294  
Db 3177 CACCCCTGTTACAAG 3190

RESULT 5  
AY405421 5069 bp DNA linear GSS 12-DEC-2003  
LOCUS Pan troglodytes HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,  
DEFINITION genomic survey sequence.  
ACCESSION AY405421  
VERSION AY405421.1 GI:39761395  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
1 (bases 1 to 5069)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
Science 302 (5652), 1960-1963 (2003)  
14671302  
2 (bases 1 to 5069)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
Location/Qualifiers  
1..5069  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>5069  
/locus\_tag="HCM2218"  
ORIGIN  
Query Match 27.3%; Score 2282.2; DB 29; Length 5069;  
Best Local Similarity 65.4%; Pred. No. 0;  
Matches 3340; Conservative 0; Mismatches 1720; Indels 45; Gaps 6;  
QY 2912 TTTGACTTGGTGACAAATGGCGGCACTCTCCATCATCCTCGGTTCCAGCGGCACCTTC 2971  
Db 2 TTTGACTTGGTGACAAATGGTGCGGCTCTCTAACTTTGGTATTGAACGATCCCCATTC 61  
QY 2972 ATCACACAGGAGCACACCCCTGTGGTGGCCATGGGATCGCTTCTTGTGATGGAACCATC 3031  
Db 62 CTCACCCAGTATCATCTGTGTGGATTCATGGAATGCTTTTATGTGATGATACCTTA 121  
QY 3032 ATCATGAGACATGAGGAGATGAGATTCCAGCTGTGACCTGAGCAATTTTGCCCGCCCC 3091  
Db 122 GTCATGAAGAAAGAGAGATGACATTCAGCTGTGATCTGATGATGATGATGATGATGAT 181  
QY 3092 AACCCAGTCTCTCTCCATCCCACTGACGCTCTTTCGCCAGCTCCTGTGAGAGAAAGGC 3151  
Db 182 AATCCCATCATGTGTCTATCACCTTTATCCACCTTTTTCAGATCTTCTCCTGAAGACAGT 241  
QY 3152 CCCATTGTGCGGAAATTCAGGCTTTGAGGAGGAAATCTCTATCTCTGGCTGCAAGATG 3211  
Db 242 CCCATCATTCGGAACACACNNNNCTCCACGAGGAACTACAAATCCAGGAACAGATTG 301  
QY 3212 AGGCTGAGCTACCTGAGCAGCGGACCCCTGGCTACAAATCTCTCCTGAGGATCAGCCTC 3271  
Db 302 AAACCTCTCCTACTTGTGATTCAGAGCTGAGGGGTATAAGTCACTTCTCAAGATCACCATG 361  
QY 3272 ACCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGTTAGGCGGTGGAGGC 3331  
Db 362 ACCAGTCTATTATTCATTAAATTAATGAAGGTTTCATCTGATGTTAGTAGGA 421  
QY 3332 CGCCTCTTCAGGAAGTGGTTCGCTGACGCCCCAGACCTGTCTTATTATTTCATTGGGAC 3391

Db 422 AGACTCTTCCAAAGTGGTTCTCGCTCACCCAACTTGGCCCTATACCTTTCATATGGGAT 481  
QY 3392 AAGACAGACGTCTACACAGAAAGGTGTTTGGGCTTTTCAGAAAGCCTTTTGTTCGTTGGGT 3451  
Db 482 AAAACAGATGCATATATAACAGAAAGTCTATGGTCTGTCTGAAGCTGTTGTCTCAGTTGGA 541  
QY 3452 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACAAACAGTGTGCAG 3511  
Db 542 TATGAGTATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 601  
QY 3512 GGCTATGAAATTGACGGCTTCCAAAGCTTGGAGGATGGAGCCCTAGACAAACATCATGCCCTC 3571  
Db 602 GGCTATGAATTGGATGGTCCCAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 661  
QY 3572 AACATTCAAAGTGGTATCCTGTCACAAAAGGGAATGGGAGAACCCAGTTTGTGTCTCAGCAG 3631  
Db 662 GATGTACAGAACGGTATACTGTACAAGGGAAACGGGGAACCAAGTTTCTCTCCAGCAG 721  
QY 3632 CCTCTGTCTATTGGGAGCATCATGGGCAATGGGCGCCGGAGAAGCATCTCCTGCCCCAGC 3691  
Db 722 CCTCCAGTCGTAGTAGCATCATGGGCAATGGGCGAAGCGCAGCAATTCCTGTGCCCCAGT 781  
QY 3692 TGC AACCGCCTTGCTGACGGCAACAAGCTCCTGGGCCCCAGTGGCCCTCACCTGTGGCTCT 3751  
Db 782 TGCAATGGTCAAGCTGATGGTAAACAAGTTACTGGGCCCCAGTGGCGCTAGCTTGTGGGATC 841  
QY 3752 GACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTCCCTCTCGGAAAT 3811  
Db 842 GACGGCAGTCTGTACGTAGGCGATTTCAACTATGTGCGGCGGATATTCCCTTCTGGAAT 901  
QY 3812 GTCACCAACATCCTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCAAGTCCAGCACAC 3871  
Db 902 GTAAACAAGTGTCTTAGAACTAAGCA-----GCAACCCAGCTCAT 940  
QY 3872 AAATACTACTGGCCACAGACCCCATGAGTGGGCGCGTCTTCCCTTTTCGACAGCAACAGC 3931  
Db 941 AGATACTACCTTGCAACGGATCCAGTACAGGAGATTTGTACGTTTCTGACACAAACACC 1000  
QY 3932 CGGCGGGTCTTTAAATCAAGTCCACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAG 3991  
Db 1001 CGCAGAATTTATCGCCCAAAGTCACTTACGGGGGCAAAAGACTTGGACTTAAATAATGCAGAA 1060  
QY 3992 GTGGTTGCGGGGACAGGTGACCACTGCTCCCTTTTGATGACACTCGCTGCGGGGATGGT 4051  
Db 1061 GTTGTGCGGGGACAGGGAGCAGTGCCTTCCGTTTGACGAGGGGAGATGTGGGGATGGA 1120  
QY 4052 GGAAGGCCACAGAACCCACACTCAACCAATCCAGGGGTATTACAGTGGACAAAGTTTGGG 4111  
Db 1121 GGAAGGCCGTGGAAGCCACACTCATGATGTCCTCCAAAGGAATGGCAGTTGATAAGAATGGA 1180  
QY 4112 CTGATCTACTTCGTGGATGGCACCATGATCAGACGCATCGATCAGAAATGGGATCATCTCC 4171  
Db 1181 TTAATCTACTTTGTTGATGGAACCATGATTAGGAAAGTTGACCCAAATGGAATCATATCA 1240  
QY 4172 ACCCTGCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGTGATTTCTGTCATG 4231  
Db 1241 ACTCTTCTGGGCTCTAACGATTTGACTTCAGCCAGACCTTTAACTTTGTGACACCAGCATG 1300  
QY 4232 GATATTTCCAGGTAAGACTGGAGTGGCCCAACAGACTTAGCCATCAACCCAAATGGACAAC 4291  
Db 1301 CACATCAGCCAGGTACGCTGGAATGGCCCACTGACCTAGCCATTAAACCTATGGATAAC 1360  
QY 4292 TCACTTTATGCTCGACAACAATGTGGTCTCGCAAAATCTCTGAAAACCAACAGGTGCGC 4351  
Db 1361 TCCATTATGCTCGGATAATAACGTAAGTTTACAGATCACTGAAAATCGTCAAGTTCCG 1420  
QY 4352 ATTGTCGCGGGAGGCCCATGCACTGCCAGTCCCTGGCATTTGACCACTTCTCTGCTAAGC 4411  
Db 1421 ATTGCTGCTGGACGGCCCATGCACTGTTCAGGTTCCCGGAGTGG---AATATCCTGTGGG 1477  
QY 4412 AAGGTGGCCATCCACGCAACCCCTGGAGTCAGCCACCGCTTTGGCTGTTTCACACAATGGG 4471

Db 1478 AAGCAGCAGTGCAGACAACTGGAATCAGCCACTGCCATTGCTGTGTCTACAGTGGG 1537  
QY 4472 GTCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCAACACT 4531  
Db 1538 GTCTGTACATTACTGAAACTGATGAGAAAGAAAATTAACCGGATAAGGCAGGTCAACA 1597  
QY 4532 AGTGGAGAGATCTCACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCC 4591  
Db 1598 GATGGAGAAATCTCCTTAGTGGCCGGAATACCTTCAGAGTGTGACTGCAAAAATGATGCC 1657  
QY 4592 AACTGTGATTTTCTTGAGACGACGATGGTTATGCAAGGATGCAAAAGTTAAATACCCCA 4651  
Db 1658 AACTGTGACTGTTACCAGAGTGGAGATGGCTACGCCAAGGATGCCAAACTCAGTGCCTCA 1717  
QY 4652 TCTTCTCTGGCTGTGTGTGCTGATGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGA 4711  
Db 1718 TCCTCCCTGGCTGCTTCTCCAGATGGTACACTGTATATTGCAGATCTAGGGAATATCCGG 1777  
QY 4712 ATTCCGGTTTATCCGGAAGAAACAAGCCTTTCCTCAACACCCAGAACATGTATGAGTGTCT 4771  
Db 1778 ATCCGGCTGTGTCAAAGGATAAGCCTTTACTTTACTTCTATGAACTTCTATGAGTTGCG 1837  
QY 4772 TCACCAATTGACCAGGAGCTCTATCTGTGTTTGATACACCCGCAAGCACCTGTACACCCAA 4831  
Db 1838 TCTCCAACTGATCAAGAACTCTACATCTTTGACATCAATGGTACTCACCAATATACTGTA 1897  
QY 4832 AGCCTGCCACAGGAGACTACCTGTACAACTTTCACCTACACTGGGGACGGCGACATCACA 4891  
Db 1898 AGTTAGTCACTGGTGATTAACCTTTTACAAATTTTACGTACAGCAATGACAAATGATATTACT 1957  
QY 4892 CTCATCACAGACAAACAATGGCAACAATGGTAAATGTCCGCCGAGACTCTACTGGGATGCC 4951  
Db 1958 GCTGTGACAGACAGCAATGGCAACACCCCTTAGAAATTAGACGGGACCCAAATCGCATGCCA 2017  
QY 4952 CTCTGGCTGGTCCAGATGGCCAGGTGTACTTGGGTGACCATGGGCACCAACAGTGCAC 5011  
Db 2018 GTTCGAGTGGTGTCTCCGATAACCAAGTGATGATGTTGACAAATAGGAACAAATGGATGT 2077  
QY 5012 CTCGAAGAGTGTGACCAACACAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCC 5071  
Db 2078 TTGAAAAGCATGACTGCTCAAGGACTGGAATTAGTTNNGTTTACTTANNNNNNNNNNNN 2137  
QY 5072 GGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACAAACATTTTATGAGTACGACAGC 5131  
Db 2138 NNN 2197  
QY 5132 TTTGGCCGCTGACAAATGTGACCTTCCCTACTGCGCCAGGTGAGCAGTTTCCGAAGTGAT 5191  
Db 2198 GAAGGTGCTGTGACAAATGTACGTTTCCAACTGAGTGTGTCAAAACCTGCATGGGGAC 2257  
QY 5192 ACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTTCAGC---AAGGATGATGTCAACATA 5248  
Db 2258 ATGACAAGGCTATCACAGTGGACATTTGAGTCTATGAGTGTAGCCGAGAGGAAGATGTACGATC 2317  
QY 5249 ACCACCAACCTGTCTGCTCAGGGCCCTTCTACACACTGTCTGCAAGACCAAGTCCGGAAC 5308  
Db 2318 ACTTCAAATCTGTCTCGATCGATTCTTTCTACACCATGGTTCAAGATCAGTTAAGAAAC 2377  
QY 5309 AGTACTACATCGGGGCGGATGGCTCCTTGGCGCTGCTGTGGCCAAACGGCATGGAGGTG 5368  
Db 2378 AGTACCAGATTGGTTATGACGGCTCCCTCAGAAATATCTACGCCAGCGGCTGGACTCA 2437  
QY 5369 GCGCTGCAGACTGAGCCCACTTGTGGTGGCTGGCACCGTCAACCCCAACCGTGGGCAAGAGG 5428  
Db 2438 CACTACCAACAGAGCGCACATTTCTGGCTGGCACTGTCTAAATCCAAACCGTTGCCAAAAGA 2497  
QY 5429 AATGTACGCTGCCATCGACAACCGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAG 5488  
Db 2498 AACATGACGTTGCTGTGGAGAACGGTCAAAACTTGGTGGAAATGGAGATTCGCGAAAAGAG 2557  
QY 5489 CAGGCTCGGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGTTGCACAAACCGAAATCTC 5548  
Db 2558 CAAGCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTC 2617

QY 5549 CTATCTCTGGACTTTGATCGCGTTAAACAGCACAGAGAAGATCTATGATGACCACCGCAAG 5608  
 Db 2618 CTTTCAGTTGACTTTGATCGAAACAACAAGACAGAAAAGATCTATGACGACCACCTTAA 2677  
 QY 5609 TTCACCCCTTCGGATTCTGTACGACCGAGCGGGCGGCCAGCCTCTGGTCACCCAGCAGC 5668  
 Db 2678 TTTCTACTGAGGATCGCCTACGACACGCTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGC 2737  
 QY 5669 AGGCTGAATGGTGTCAACGTGACATACCTCCCTGGGGGTACATTTGCTGGCATCCAGAGG 5728  
 Db 2738 AAGCTGATGGCGTCAATGTCAACCTATTATCCACAGGTCAAAATGGCAGCATCCAGCGA 2797  
 QY 5729 GGCATCATGTCTGAAAGAAATGGATACGACACCGCGGCCGATCATCATCCAGGATCTTC 5788  
 Db 2798 GGCACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTT 2857  
 QY 5789 GCTGATGGGAAGACATGGAGCTACACATACCTTAGAGAAGTCCATGGTGTCTGCTACTACAC 5848  
 Db 2858 GCTGATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCAT 2917  
 QY 5849 AGCCAGAGGAGTATATCTTTGAGTTCGACAAGAAATGACCGCTCTCTCTGAGCATG 5908  
 Db 2918 AGCCAGCGGAGTACATCTTCGATACGATATGTGGACCGCTGTCTGCCATCACCATG 2977  
 QY 5909 CCCAACGTGGCGGCGAGACACTAGAGACCATCCGCTCAGTGGGTACTACTACAGAAACATC 5968  
 Db 2978 CCCAGTGTGGTGGCCACACCAATGCAGACCATCCGATCCATGGCTACTACCGCAACATA 3037  
 QY 5969 TATCAGCCCTGAGGGCAATGCCTCAGTACATACAGGACTTCACCTGAGGATGGGCACTC 6028  
 Db 3038 TACAACCCCGGAAAGCAACGCTCCATCATCACGGACTACAAACGAGGAGGGCTGCTT 3097  
 QY 6029 CTTTACACCTTCTACCTGGGCACTGGCGCAGGTTGATATACAAAGTATGGCAAACTGTCA 6088  
 Db 3098 CTGCAAAACAGTTTCTTGGTACAAGTCGGAGGTCTTATTCAAATACAGAAAGGCACT 3157  
 QY 6089 AAGCTGGCAGACGCTCTATGACACCAACCAAGGTCAGTTTACCTATGACGAGACGGCA 6148  
 Db 3158 AGGCTCTCAGAAATTTTATATGATGACGAGAGTCAGTTTACCTACGATGAACAGCA 3217  
 QY 6149 GGCATGCTGAACCATCAACCTACAGAAATGAGGGCTTCACCTGCACCATCCGCTACCGT 6208  
 Db 3218 GGAGTCCCTAAAGACAGTAAACCTCCAGAGTGACGGTTTATTGACCATTAGATACAGG 3277  
 QY 6209 CAGATTGGGCCCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAAC 6268  
 Db 3278 CAAATTGGTCCCTGATTGACAGGCAGATTTTCCGCTTTTAGTGAAGATGGGTAAAT 3337  
 QY 6269 GCCCGTTTTGACTAAACTATGACAAACAGCTTCCGGGTGACGAGCATGCAAGCTGTGATC 6328  
 Db 3338 GCAAGATTTGACTATAGCTATGACAAACAGCTTTCGAGTGACAGCATGCAAGGTGTGATC 3397  
 QY 6329 AACGAGACCCCACTGCCCATTGATCTCTATCGCTATGATGATGTGTGAGCAAGACAGAG 6388  
 Db 3398 AATGAAACTCCCACTGCCTATTGATCTGTATCAGTTTGTATGACATTTCTGGCAAGTTGAG 3457  
 QY 6389 CAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACAGATCATCACCAAGTGTGTC 6448  
 Db 3458 CAGTTTGGAAAGTTTGGAGTTATATATATGATATTAACAGATCATTTCTACAGCTGTN 3517  
 QY 6449 ATGACCCACACCAAGCATTTTGTATGATATGGCAGGATGAAGAAAGTGCAGTATGATC 6508  
 Db 3518 ATGACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGAGAGATTCAATATGAGATA 3577  
 QY 6509 TTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATATGATGAGGCGGAGTAGTGAAG 6568  
 Db 3578 TTCAGGTGCTCATGTACTGGATTACAATTCAGTATGATGATGATGATGATGATGATGATG 3637  
 QY 6569 AAGGAGCTGAAGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGAGTATGATGCT 6628  
 Db 3638 AGAGAGATTAAATAGGGCCCTTTGCCAACACCAACCAATATGCTTATGAATATGATGTT 3697

QY 6629 GACGGCCAGCTGCAGACAGTCTTCCATCAATGACAAGCCCACTCTGGCGCTACAGCTACGAC 6688  
 Db 3698 GATGGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACATACGAT 3757  
 QY 6689 CTCAATGGAAACCTTGCACTTACTGAGCCCTGGGAAACAGTGCACGGCTCACACCACTACGG 6748  
 Db 3758 CTGAATGGAAACCTCCATTTACTGAAACCCAAAGTAAACAGTGGCGCTGTGACACCCCTTCGC 3817  
 QY 6749 TATGACATCCCGCAGACCGCATACTCGCTGGGTGACACGTGCAATAACAAGATGGATGAGGAT 6808  
 Db 3818 TATGACCTGCGAGACAGAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGAT 3877  
 QY 6809 GGCTTCTGAGGAGCGGGGCTGGAGTGTAGTATCTTTGAGTACAACTCAGCTGGCCTGCTCATC 6868  
 Db 3878 GGTTTCTCTACGTCAAAGGGGCAAGGAAATCTTTGAATATAGCTCAAAGGGCTTCTAACT 3937  
 QY 6869 AAGGCCTACAACCGGGCTGGCAGTGGAGTGTAGGTACCGCTACGATGGCCTGGGCGG 6928  
 Db 3938 CGAGTTTACAGTAAAGGCGAGTGGTGACAGTGANNACCGTTATGACGGCTGGGAAGG 3997  
 QY 6929 CGCGTGTCCAGCAAGAGCAGCCACAGCCACCACTCCAGCTCTGAGATCACCTCCCTCTAC 6988  
 Db 3998 CGTGTNNNAGCAAAACCACTGTCAGGACAGCACCTGCACTTTTATGCTGACTTAACT 4057  
 QY 6989 AACCCCAACCAAGGTCAACCCACCTGTACAACCACTCCAGCTCTGAGATCACCTCCCTCTAC 7048  
 Db 4058 TATCCCACTAGGATTACTCATGTCTACAACCATTCGAGTTCAGAAATTAACCTCCCTGTAN 4117  
 QY 7049 TACGACTTGAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGTGATGAGTTTACATA 7108  
 Db 4118 NNTGATCTCAAGGACATCTTTTGGCATGGAAATCAGCAGTGGGATGAATTTCTATAT 4177  
 QY 7109 GCTTGTGCAACATCGGGACCCCTCTTGTCTCTTTTAGTGAACAGGTTTGTGATGATCAAG 7168  
 Db 4178 GCATCAGATAACACAGGGACACCACTGGCTGTGTTTCAGTAGCAATGGCTTATGCTGAAA 4237  
 QY 7169 CAAATCCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCACTTTTCAGATC 7228  
 Db 4238 CAGATTCACTACCTGATATGTTGAAATCTATTTTTCAGTCTAATATTGACTTTTCACTG 4297  
 QY 7229 ATCATAGGCTACCATGGTGGCTCTATGATCCACTCACCAAGCTTTGTCACATGGGCGCG 7288  
 Db 4298 GTAATTGGATTTTCATGGTGGCTGTATGACCCACTCACCAAAATTAATCCACTNNNNNN 4357  
 QY 7289 CGAGATTATGATGTGTGGCCGGACGCTGGAGTACGCCAGACCAAGCTGTGGAAGCAC 7348  
 Db 4358 NNN 4414  
 QY 7349 CTTAGTAGCAGCAACGTCATGCTTTTAACTCTCTATATGTTCAAAAACAAACCCCATC 7408  
 Db 4415 AGAATCGGGAAGGACCCAGCTCCTTTTAACTTTGATGATGTTAGGAATAACAACCCCTGCA 4474  
 QY 7409 AGCAACTCCAGGACATCAAGTGTCTCATGACAGATGTTAAACAGCTGGTGTCACTTTT 7468  
 Db 4475 AGCAAAATCCATGACGTGAAGATTACATCACAGATGTTAAACAGCTGGTGTGACATTT 4534  
 QY 7469 GGATTTCCAGTACACAACGTCATCCCTGGTTATCCCAAAACCCAGACATGGATGCCATGAA 7528  
 Db 4535 GGTTTCCATCTGCACATGCTATTCTCTGGATTCCCTGTTCCCAAAATTTGATTAAACAGAA 4594  
 QY 7529 CCCTCCTACGAGCTCATCCACACACAGATGAAAACGAGAGTGGGACCAACAGCAAGTCT 7588  
 Db 4595 CCTTCTTACGA-----ACTTGTGAAGATCAGCAGTGGGATGATGTACCGCCC 4642  
 QY 7589 ATCCTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCTTTGTCACTTAGAACGG 7648  
 Db 4643 ATCTTCGGAGTCCAGCAGCAAGTGGCGCGGAGGCGCAAGGCTTCTGTGCTGGGGAAG 4702  
 QY 7649 TTTGACCAAGCTCTATGGCTCCACATCACAGCTGCCAGGAGTCCAAAGACCAAGAAAG 7708  
 Db 4703 ATGGCCGAGGTGAGGTGAGCCGCGCGCGCGCGC---GGCGGCGAGTCTCTGCTGTGG 4759  
 QY 7709 TTTGCATCCAGCGGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGA 7768

Db	4760	TTGCCACGGTCAAGTCGCTGATCGGCAAGGGCGTCAATGCTGGCCGTCAGCCAGGGCCGC	4819
Qy	7769	GTGACCACAGACATCATCAGTGTGGCCAAATGAGGATGGGCGAAGGGTTGCTGCCATCTTG	7828
Db	4820	GTGAGACCAACGTCCTCAACATCGCCAACGAGGACTGCATCAAGGTGGCGCGTGTCTC	4879
Qy	7829	AACCATGCCCACTACCTAGAGAACTGCACCTTCAACATGATGGGTGGATACCCATTAC	7888
Db	4880	AACAACGCCTTCTACCTGGAGAACCTGCACCTTCAACATCGAGGGCAAGGACACGCATAC	4939
Qy	7889	TTTGTGAAACAGGACCTTCAAGAGGTGACCTGGGCATCTGGGCCTCAGTGGGGGGCGG	7948
Db	4940	TTTCATCAAGACCACACCGCCGAGAGCGACCTGGACGCTGGGCTGACGAGCGGCCGC	4999
Qy	7949	CGAACCCCTGGAGAAATGGGTTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAATGGC	8008
Db	5000	AAGCGCTGGAGAACGGCATCAACGTCACGGTGTCCGAGTCCACCACCGTGGTGAACGGC	5059
Qy	8009	AGGAC 8013	
Db	5060	AGGAC 5064	
RESULT 6			
AY405422			
LOCUS			
DEFINITION Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.			
ACCESSION AY405422			
VERSION AY405422.1 GI:39761396			
KEYWORDS GSS.			
SOURCE Mus musculus (house mouse)			
ORGANISM Mus musculus			
REFERENCE 1 (bases 1 to 5087)			
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios			
JOURNAL Science 302 (5652), 1960-1963 (2003)			
PUBMED 14671302			
REFERENCE 2 (bases 1 to 5087)			
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.			
TITLE Direct Submission			
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA			
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.			
FEATURES			
source			
1..5087			
/organism="Mus musculus"			
/mol_type="genomic DNA"			
/db_xref="taxon:10090"			
<1..>5087			
/locus_tag="HCM2218"			
ORIGIN			
Query Match 26.7%; Score 2232.8; DB 29; Length 5087;			
Best Local Similarity 68.4%; Pred. No. 0;			
Matches 3178; Conservative 0; Mismatches 1432; Indels 36; Gaps 5;			
Qy	2912	TTTGACTTGGTGACAAATGGCGGCACTCTCCATCATCTCGGTTTCGAGCGGCACCTTTC	2971
Db	1	TTTGACTTGGTGACAAATGGTGCGCTTCTCTGACTTTGGTATTTGAGCGTTCCCCATT	60
Qy	2972	ATCACACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTTCTTTGTCATGGAACCATC	3031

Db	61	CTCACTCAGTACCACACTGTGTGGATTCCCTGGAATGCTTTTATGTGATGATACCCCT	120
Qy	3032	ATCATGAGACATGAGGAGATGAGATTCCCAGCTGTGACCTGAGCAATTTTGCCCGCCCC	3091
Db	121	GTCTGAAGAAAGAGGAGAACGACATTCCAGCTGTGACCTCAGTGGCTTTGTGAGGCCA	180
Qy	3092	AACCCAGTCGTCTCTCCATCCCACTGACGTCTTTCGCGAGCTCCTGTGCAGAGAAAGGC	3151
Db	181	AGTCCCATCATTTGTGTCTTCCCGTTATCCACCTTCTTCAGGTCTTCCCTGAGGACAGC	240
Qy	3152	CCCATTTGCCCCGAAATTCAGGCTTTGCAGGAGGAAATCTCTATCTCTGGCTGCAAGATG	3211
Db	241	CCCATCATCCCCGAGACACAGGTCCTGCATGAAGAAACCAATTCAGGAACAGATTG	300
Qy	3212	AGGCTGAGCTACCTGAGCAGCGGACCCCTGGCTACAAATCTGTCTCTGAGGATCAGCCTC	3271
Db	301	AAACTTTCCTACCTGAGTTCAGAGCGGCAGGGTACAGTCAGTCTTTAAGATTACCATG	360
Qy	3272	ACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGGTAGCGGTGAGGGC	3331
Db	361	ACCCAGGCGTCATACCGTTTAACTCATGAAGTCCATCTGATGGTGGCCGTGTTGGG	420
Qy	3332	CGCTCTTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTCTATTATTTTCATTTGGGAC	3391
Db	421	AGACTCTTCCAGAAAGTGGTTTCTCGCTCGCCAACTTGGCCTACACGTTTCATCTGGAT	480
Qy	3392	AAGACAGACGTCCTAACACAGAAAGTGTTTGGGCTTTCAGAAAGCCTTTGTTTCCGTGGT	3451
Db	481	AAGACGGACGCATATAATCAGAAAGTCTACGGCTTGTGAGAGGCAGTTGTGTCCGTCCGA	540
Qy	3452	TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAAACAACAGTGTCTGCAG	3511
Db	541	TACGAGTACGAGTCGTGCTTGGACCTGACTCTCTGGGAAAAAGAGGACTGCCGTTTGGAA	600
Qy	3512	GGTATGAAATTGACGCGTCCAAGCTTGGAGGATGGAGCCTTAGACAAACATCATGCCTC	3571
Db	601	GGCTATGAGTTGGATGCTTCGAAACATGGCGGCTGGACGTTGGACAAGCACCATGTACTG	660
Qy	3572	AACATTCAAAGTGGTATCCTCTGCACAAAGGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG	3631
Db	661	GACGTTCAGAACGGTATACTATACAAAGGAAATGGAGAAATCAGTTTCATCTCTCAGCAG	720
Qy	3632	CCTCCTGTCTATGGGAGCATCATGGGCAATGGGCGCCGAGAGCATCTCCTGCCCCAGC	3691
Db	721	CCTCCGGTGGTCAGCAGCATCATGGGTAATGGTCGGAGGCGTAGCATCTCATGCCCAAGT	780
Qy	3692	TGCAACGGCCTTGTCTGACGSCAACAAAGCTCCTGCGCCCCAGTGGCCCTCACCTGTGGTCT	3751
Db	781	TGCAATGGTCAAGCTGACGGGAACAACCTCCTGGCACCCGCTGGCGCTTGCCCTGTGGGATC	840
Qy	3752	GACGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTCCCCTCTGGAAT	3811
Db	841	GACGGCAGTCTATACGTAGGGGATTTCAATTACGTCGGCGGATATTCCCGTCTGGGAAT	900
Qy	3812	GTCAACCAACATCCTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCACAGTCCAGCACAC	3871
Db	901	GTGACAAAGTGTTTTAGA-----ACTAAGTAGCAACCCAGCTCAC	939
Qy	3872	AAATACTACCTGGCCACAGACCCCATGAGTGGGGCCGTCTTCTTTCTGACAGCAACAGC	3931
Db	940	AGATACTACCTGGCTACGGACCCAGTCAACGGAGATTTGTACGTCTCTGTATACTAACACC	999
Qy	3932	CGCGGGGTCTTTAAATCAAGTCCACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAG	3991
Db	1000	CGCAGAATCTATCGGCCGAAATCACTACGGGAGGCCAAAGACCTGACTAAAAACGCTGAA	1059
Qy	3992	GTGGTTCGGGGACAGGTACACAGTGCCTCCCTTTCGATGACACTCGCTCGGGGATGGT	4051
Db	1060	GTGGTTCGGGACCCGGGAACAGTGCCTTCCCTTTGACGAGGCCAGGTGTGGGATGGA	1119
Qy	4052	GGGAAGGCCACAGAACCCACACTCACCAATCCCAGGGGTATTACAGTGGACAAGTTGGG	4111

Db 1120 GGCAAGGCTGTGGAAGCAACGGCTCATGAGTCCCAAGGAATGGCAATCGATAAGAACGGA 1179  
QY 4112 CTGATCTACTTCGTGGATGGCACCATGATCAGACGCGATCGATCAGAAATGGGATCATCTCC 4171  
Db 1180 CTGATCTACTTTGTGTGTAAGAACCATGATCAGAAAGGTTGATCAAAATGGAATCATATCA 1239  
QY 4172 ACCCTGCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGTGATTCTGTGATG 4231  
Db 1240 ACTCTCCTGGGCTCCAACGACCTCACGTCAGCTCGACCTTTAACTGTGATACTAGCATG 1299  
QY 4232 GATATTTCCAGGTAAGACTGGAGTGGGCCACAGACTTAGCCATCAACCCCAATGGACAAC 4291  
Db 1300 CATATCAGCCAG-----CTGGAATGGCCCACTGACCTCGCGATCAACCCCATGGATAAC 1353  
QY 4292 TCACCTTTATGTCCTCGACAACAATGTGGTCCTGCAAAATCTCTGAAAACCAACAGGTGGC 4351  
Db 1354 TCCATCTACGTCCTGGATAATAACGTAAGTTTACAGATCACTGAAAACCGTCAGGTCCGC 1413  
QY 4352 ATTGTGCGCGGAGGCCCATGCACTGCCAGGTCCCTGGCATTGACCACTTCTCTGTAAAGC 4411  
Db 1414 ATCGTGC CGGCGGCCCATGCACTGTCAAGTCCCTGGAGTGG---AATACCGGTGGGG 1470  
QY 4412 AAGGTGGCCATCCACGCAACCCCTGGAGTCAGCCACCGCTTTGGGTGTTTCAACAAATGGG 4471  
Db 1471 AAGCACGCGGTTCAGACCAACCCCTGGAGTCAGCCACGGCCATTGCTGTCTCTACAGCGGG 1530  
QY 4472 GTCCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCAACCACT 4531  
Db 1531 GTCCTTTTACATACCGGAAACTGATGAGAAAGATCAACCGAATAAGGCAGGTCAAGACA 1590  
QY 4532 AGTGAGAGATCTCACTCGTTGCTGGGCGCCCGAGTGGCTGTGACTGTAAAAATGATGCC 4591  
Db 1591 GACGGGAGATCTCCTTAGTGGCTGGGATACCTTCGGAATGTGACTGCAAGAACGACGCC 1650  
QY 4592 AACTGTGATTGTTTCTTGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCA 4651  
Db 1651 AACTGTGACTCTACCAAGCGGAGACGCTACGCCAAAGATGCCAAACTCAATGCGCGG 1710  
QY 4652 TCTTCCCTTGGCTGTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAACATCCGA 4711  
Db 1711 TCCTCCCTGGCGCCTCGCCAGATGGCACTCTGTACATTGCGAGATCTGGGAAATATCAGG 1770  
QY 4712 ATTCGGTTTATCCGGAAGAACAAAGCCTTTCCTCAACACCCAGAACATGTATGAGCTGTCT 4771  
Db 1771 ATCCGGGCGGTTTCGAAGAAATAAACCTTACTGAACTCAATGAACCTTTACGAAGTTGCC 1830  
QY 4772 TCACCAATTGACAGGAGCTCTATCTGTTGTATACCAACCGGCAAGCACCTGTACACCCAA 4831  
Db 1831 TCTCCAACTGATCAAGAGCTCTACATCTTTGACATCAACGGTACTCACCAGTACACCGTG 1890  
QY 4832 AGCCTGCCACAGGAGACTACCTGTACAACTTACCTACACTGGGGACGGCGACATACACA 4891  
Db 1891 AGCCTGGTCACGGGTGACTACCTATATAATTTTAGTTACAGCAATGACAAATGACGTCAAC 1950  
QY 4892 CTCAATCAGACAACAATAATGGCAACATGGTAAATGTCCGCCGAGACTCTACTGGGATGCC 4951  
Db 1951 GCTGTAACGTACAGCAATGGCAACACCCCTCCGAATCCGAAGGGATCCGAATCGGATGCCG 2010  
QY 4952 CTCTGGTGGTGGTCCCAGATGGCCAGGTGTAATCTGGGTGACCATGGGCACCAACAGTGCA 5011  
Db 2011 GTCGGGTGGTGTCTCCTGTATAACCAAGTGATATGTTGACCATAGGCACCAACGGGTGT 2070  
QY 5012 CTCAAGAGTGTGACCAACACAAGGACAGAGTTGGCCATGATGACATACCATGGCAATTC 5071  
Db 2071 CTGAAAAGCATGACCGCTCAGGGCCTGGAACCTGTTTGTGTTACTTACCATGGCAACAGT 2130  
QY 5072 GGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACAACATTTTATGATACGACAGC 5131  
Db 2131 GGGCTTTTAGCCACCAAAAGTGACGAAACTGGATGGACAACATTTTGTGACTATGACAGT 2190  
QY 5132 TTTTGGCGCCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCCGAAGTGAT 5191  
Db 2191 GAAGGTGCGCCTGACGAATGTTACCTTCCCACTGGGGTGGTTACAAACCTGCACCGGGAC 2250

QY 5192 ACAGACAGTTTCACTAGTGCATGTCCAGGTAGAGACCTCCAGCA---AGGATGATGTCAACATA 5248  
Db 2251 ATGGACAAGGCTATCACGGTGGACATCGAGTCAATCCAGCAGAGAGGAAGATGTACGATC 2310  
QY 5249 ACCACCAACCTGTCTGCCCTCAGGGCCTTCTACACACTGTGTCGAAGACCAAGTCCGGAAC 5308  
Db 2311 ACTTCGAACCTGTCTCCATCGATTCCTTCTACACCAATGGTCCAAGACCAAGTTAAGAAAC 2370  
QY 5309 AGCTACTACATCGGGGCGGATGGTCCCTTGGCTGCGGTGCTGTGGCCCAACGGCATGGAGTG 5368  
Db 2371 AGTTACCAGATTGGGTATGATGGCTCCCTTAGAATCTTCTATGCCAGTGGTCTGGACTCT 2430  
QY 5369 GCGCTGCAGACTGAGCCCCACCTTGTGGCTGGCACCCTGTCACCCCTGCGGCAAGAGG 5428  
Db 2431 CACTACCAGACAGAGCCCCACGTTCTGGCTGGCAGCGGAATCCACAGTAGCCAAAAGA 2490  
QY 5429 AATGTCACGTGCCTCCATCGACAACCGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAG 5488  
Db 2491 AACATGACTCTTCCCGGTGAGAACGGGCAGAAATCTGGTGGAGTGGAGATTCCGAAAAGAA 2550  
QY 5489 CAGGCTCGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGCTGCACAAACCGAAATCTC 5548  
Db 2551 CAAGCCCAGGSCAAAGTCAACGTAATTCGGCCGGAAGCTCAGGTCATGCGGCGCAACCTA 2610  
QY 5549 CTATCTCTGGACTTTGATCGCGTAAACGCAACAGAGAAAGATCTATGATGACCCAGCAAG 5608  
Db 2611 CTCTCAGTGGACTTTGATCGGACCAACCAAGACGGAAGATCTATGATGACCCAGCGGAAA 2670  
QY 5609 TTCACCCCTTCGGATTCTGTACGACGAGCGGGCGGCCAGCCTCTGGTCACCCAGCAGC 5668  
Db 2671 TTTCTCCTGAGGATCGCTTACGACACGCTCGGGCACCCGACTCTCTGGCTGCCGAGTAGC 2730  
QY 5669 AGCTGAATGGTGTCAACGTGACATACTCCCCTGGGGGTTACATTTGCTGGCATCCAGAGG 5728  
Db 2731 AAGCTAATGGCAGTGAACGTCACTACTCATCCACCGGTCAAAATGCCAGCATCCAGAGA 2790  
QY 5729 GGCATCATGTCTGAAAAGAAATGGAATACGACCGGCGGCCGATCACATCCAGGATCTTC 5788  
Db 2791 GGGACCAACGAGCGAAAAGGTGGACTATGACAGCCAGGGGAGGATCGTATCTCGGGTCTTT 2850  
QY 5789 GCTGATGGGAAGACATGGAGCTACACATACTTAGAGAAGTCCATGGTGTCTACTACAC 5848  
Db 2851 GCCGATGGGAAAACATGGAGTTACACGTACTTGGAAAAGTCCATGGTTCTTCTGCTCCAT 2910  
QY 5849 AGCCAGAGGCGAGTATATCTTTGAGTTCGACAAGAAATGACCGCCTCTCTTCTGTGACGATG 5908  
Db 2911 AGCCAGCGGCGAGTACATCTTCGAATACGACATGTGGACCGCCTGTCCGCCATCACCATG 2970  
QY 5909 CCCAACGTGGCGGCGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAAACATC 5968  
Db 2971 CCCAGTGTGGCTCGCCACACCATGACACCATCCGCTCCATTTGGCTACTACCGCAACATC 3030  
QY 5969 TATCAGCCCTCTGAGGGCAATGCTCAGTCAATACAGGACTTCACTGAGGATGGGCACCTC 6028  
Db 3031 TACAATCCCCCAGAAAAGCAATGCCTCTATCATCACCGACTACAACGAGGAAGGCTGCTT 3090  
QY 6029 CTTACACACCTTCTACCTGGGCACTGGCCGCGAGGGTGATATAACAAGTATGGCAAACTGTCA 6088  
Db 3091 CTGCAAAACAGCTTTCTTGGAAACGAGTCGGAGGGTCTTATTCAAGTATAGAAGGCAGACC 3150  
QY 6089 AAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAAGTTTCACTATGACGAGACGCA 6148  
Db 3151 AGGCTATCAGAAATTTTATACGACAGCACAAGAGTCAGTTTTTACCTACGACGAAACAGCG 3210  
QY 6149 GGCATGTGAAGACCATCAACCTACAGAATGAGGGCTTCACTGCACCATCCGCTACCGT 6208  
Db 3211 GGAGTCTGAAAACAGTAAACCTTCAGAGTGTGTTTATTTGACCATTAGATACAGG 3270  
QY 6209 CAGATTGGGCCCTGTATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAAC 6268  
Db 3271 CAAATTGGTCCCTGTATTGACAGACAGATTTTCCGCTTTCGCGAGGATGGAATGGTAAAT 3330



Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,  
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
CDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/  
Location/Qualifiers

FEATURES  
source

1. .3038  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:A130060L24"  
/db\_xref="MGI:2402337"  
/db\_xref="taxon:10090"  
/clone="A130060L24"  
/tissue\_type="thymus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="16 days neonate"  
<1. .2978  
/notes="unnamed protein product; odd Oz/ten-m homolog 1  
(Drosophila) (MGD|MGI:1345185, GB|NM\_011855, evidence:  
BLASTN, 99%, match=3037)  
putative"  
/codon\_start=3  
/protein\_id="BAC29893.1"  
/db\_xref="GI:26332326"  
/translation="EINLSSEPHILAGAVNPTLKGKINISLPGEHNANLIEWRQKEQN  
KGNVSAFERLRNHRNLLSIDFDHMTRTGKIYDHRKFTLRILYDQTRFILWSPVS  
RYNEVNITYSPGLVTFIQRGTWNEKMEYDQSGKIISRTWADGKIWSYTVLEKSVMLL  
LHSORRYIFEYDQSDCLLSVTMPMSVRHSLQTMLSVGYRNIYTPPDSSTFPIQDYSR  
DGRLLQTLHLGTGRRVLYKYTKQARLSEILYDTQVLTLYEESGVIKTHLMHDGFI  
CTIRYQTGPLIGRQIFRFESEGLVNARFDYSYNNFRVTSQAVINETPLFDLYRYV  
DVSGRTEQFGKFSVINYDLNQVITTTVMKTKIFNANGQVIEQYELKAIAYWMTIQ  
YDNMRMVICDIRVGDANITRYFEYDADGQLQTVSVNDKIOWRYSYDLNGNINLLS  
HGNSARLTPLRYDLRDRITLGEIQYKDEDEGFLRQGNIDIFEYNSNGLIQKAYNKVS  
GWTQVYYDGLGRRVASKSLGQHLQFFYADLANPRTVTHLNYHTSAEITSLYYDLQ  
HLIAMELSSGEYVACDNMGTPLAVFSSRGQVTKBILYTPYGDYHDFYDPFEVIG  
FHGGLYDFLTKLVHLGQRDYDVAGRWTTTPNHHIWKQLNLLPKPFNLYSFNNYPVGK  
IQDYAKYTTDGTWLELFGQLHNVLPGPFPKPELNMELTYELLQLOTKTQEWDPGKM  
ILGICQELQKQLRNFIQLDLPMTPOYNNEGRCLGKQPRFAAVPSVFGKIKFAIKE  
GIVTADIIGVANEDSRRLAAILNNAHYLENLHFTIEGRDTHYFIKLGSLBEDLVLI  
TGGRILENGVNVTVSMTSVLNGRTRRFADIQHGLALCFNIRYGTTVBEEKNHVLE  
MARORAVAQAWTOEQRLOGBEETRVTWTEGEKQQLLGTRVQGYDGYFVLSVEQYLE  
LSDSANNIHFMRQSEIGR"

ORIGIN

Query Match 16.9%; Score 1407.8; DB 11; Length 3038;  
Best Local Similarity 67.4%; Pred. No. 8.9e-276;  
Matches 2014; Conservative 0; Mismatches 967; Indels 6; Gaps 2;  
QY 5362 GGAGGTGGCGTGCAGACTGAGCCCACTTGTCTGGTGGCAGCGGTCAACCCACCGTGGG 5421  
DDB 2 GGAGATCAACCTCAGCTCGGAGCCCCACATCCTGGCAGGGGTGTCAACCCCACTTAGG 61  
QY 5422 CAAGAGGAATGTCAGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCG 5481

Db 62 CAAATGCAACATATCGTTGCTGAGAGCACAATGCAAACTCATTGAGTGAGACAGAG 121  
QY 5482 CAAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCGCGCGGTGCGGGTGCAACCCG 5541  
Db 122 AAAGGAGCAAAACAAAGGCAACGTTTCAGCTTTTGAGAGAAGACTGAGGGCCCAATAACAG 181  
QY 5542 AAATCTCCTATCTCTGGACTTTTGATCGCGTAAACACGCACAGAGAAGATCTATGATGACCA 5601  
Db 182 AAACCTACTTTCCATTTGATTTGATCATATGACCCGCACGGGTAAGATCTATGATGACCA 241  
QY 5602 CCGCAAGTTCAACCTTCGGATTCTGTACGACCAAGCGGGGGCGGCCAGCCTCTGGTCAAC 5661  
Db 242 CCGAAAATTCACCTCTTCGAATTCCTTTATGACCAAGACCGGGCGACCTATTCTGTGGTCTCC 301  
QY 5662 CAGCAGCAGGCTGAATGGTGTCAACGTGACATACCTCCCTGGGGTTACATTTGCTGGCAT 5721  
Db 302 TGTGAGCAGATATAATGAAGTGAACATCAGTATTCACTTCAGGATTAGTCACATTTAT 361  
QY 5722 CCAGAGGGGCATCATGTCTGAAAGATGGAATACGACCAAGCGGGGGCGGCATCACATCCAG 5781  
Db 362 TCAAAGAGGACCTGGAAATGAAAGATGGAGTATGACCAAGATGCAAAATAATCTCAAG 421  
QY 5782 GATCTTCGCTGATGGGAAGACATGGAGCTACACATACCTTAGAGAAGTCCATGGTGTCT 5841  
Db 422 AACTTGGGCCGATGGGAAAATTTGGAGCTATACCTACTTAGAAAAGTCTGTGATGCTTCT 481  
QY 5842 ACTACACAGCCAGAGGCGAGTATATCTTTGAGTTTCGACAAGATGACCGCCTCTCTTCTGT 5901  
Db 482 GTTACACAGTCAGCGGCGTTTACATCTTTGAGTATGACCAATCAGATTTGCCTGCTCAGT 541  
QY 5902 GACGATGCCCAACGTGGCGCGGACACATAGAGACCATCCGCTCAGTGGGCTACTACAG 5961  
Db 542 TACCATGCTAGCATGTGCGCCACAGCTTACAAACCATGCTTTTCACTGGGCTACTACCG 601  
QY 5962 AAACATCTATCAGCCCCCTGAGGGCAATGCTCAGTCATACAGGACTTCACTGAGGATGG 6021  
Db 602 TAATATCTACACCCCAACAGACAGACGACGACCTCTTTTATCCAAGACTACAGTCGAGATGG 661  
QY 6022 GCACCTCCTTACACACCTTCTACCTGGGCACTGGCCGAGGGTGATATACAAGTATGGCAA 6081  
Db 662 CCGACTGTTACAGACTCTGCATCTAGGACTGGGCGCAGAGTTTATACAAGTACACAAA 721  
QY 6082 ACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTGAGTTTCACTATGACGA 6141  
Db 722 GCAAGCAAGACTGTCTGAGATTCTCTATGACACCACTCAGGTACACATTAAACATATGAAGA 781  
QY 6142 GACGGCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGSGTTTCACTGCACCATCCG 6201  
Db 782 GTCTTCTGGAGTAATTAGACAATACACCTGATGATGACGGAATTCATCTGCACAATCAG 841  
QY 6202 CTACCGTCAGATTGGGCCCCCTGATGACCGACAGATCTTCCGCTTCACTGAGGAAGGCAT 6261  
Db 842 ATACAGGCAACAGGACCTCTCATTTGGACGTGAGATTTTCCGATTCAGTGAAGAAGGTCT 901  
QY 6262 GGTCAACGCGCGTTTGTACTACAACATGACAACAGCTTCCGGGTGACCAAGCATGCAGGC 6321  
Db 902 TGTGAATGACGGTTTGTACTACA---GTTACAACAACCTTTCTGTGTCAACGAGCATGCAAGC 958  
QY 6322 TGTGATCAACGAGACCCCACTGCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAA 6381  
Db 959 TGTGATCAATGAGACCCCTTTGCCCATAGATCTGTACCGATATGTTTCTGTGTAG 1018  
QY 6382 GACAGAGCAGTTTGGGAAGTTTGTGTCTATTACTATGACATTAACCAAGATCATCAACAC 6441  
Db 1019 AACAGAACAGTTTGGAAAATTCACTGTAATTAATTATGATTTAAATCAGGTTTATACTAC 1078  
QY 6442 AGCTGTATGACCCACACCAAGCAATTTGATGATATGAGGAGTGAAGGAAGTGCAGTA 6501  
Db 1079 AACAGTGATGAAGCACACCAAGATTTCAATGCCAATGGCAAGTCAATGAGGTCCAATA 1138  
QY 6502 TGAGATCTTCGCTCGTCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGT 6561  
Db 1139 TGAGATCCTAAAGGCAATTGCTTACTGGATGACTATTTCAGTATGATAATATGGGCCGAT 1198

QY	6562	AGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACCTCGCTACTCTCTATGAGTA	6622
DB	1199	GGTCATATGTGATATAAGGGTAGGAGTAGATGCCAATATCACAAGGTACTTCTATGAATA	1258
QY	6622	TGATGCTGACGSCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGCGCTACAG	6681
DB	1259	TGATGCTGATGSGCAACTTCAAACTGTTCTGTGAATGACAAAATCCAGTGGCGGTATAG	1318
QY	6682	CTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGSGCTCACACC	6741
DB	1319	TTATGATCTGAATGGAACATCAACCTTTTGAGCCATGGGAATAGTGCTCGTCTTACTCC	1378
QY	6742	ACTACGGTATGACATCCGACCGCATCACTCGCTGGGTGAGTGCATGCAATACAAGATGGA	6801
DB	1379	TCTCCGATATGACCTCCGAGACCGCATCACTAGACTAGGGGAATTCAGTATAAAATGGA	1438
QY	6802	TGAGGATGGCTTCTTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCCT	6861
DB	1439	TGAAGACGGTTTTCTTGAGACAGAGGGGAATGATATATTTGAGTATAATTTCTAATGGTCT	1498
QY	6862	GCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCCT	6921
DB	1499	GCTGCAGAAAGCCTACAATAAAGTTTCTGGCTGGACTGTACAGTATTTACTATGATGGGCT	1558
QY	6922	GGGGCGCGCTGTCCAGCAAGAGCAGCCACAGCCACCCTGCAGTTCGAGTACCGCTACGATGCAGA	6981
DB	1559	TGGCGACGCGTTGCCAGTAAGTCCAGCCTAGGACAGCATCTCCAATTTCTTTATGCAGA	1618
QY	6982	CCTGACCAACCCACCAAGGTCAACCCCTGTACAACCACTCCAGCTCTGAGATCACCTC	7041
DB	1619	CCTGGCCAAACCTATAAGAGTTACTCATCTGTACAACCAACAAGCGCAGAGATTACATC	1678
QY	7042	CCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTCAATGAT	7101
DB	1679	TCCTTATTATGATCTCCAAGGTCACCTTATTGCCATGGAGCTAAGCAGTGGTGAGGAATA	1738
QY	7102	TTACATAGCTTTGTGACAAACATCGGACCCCTCTTGTGTCTTTTAGTGAACAGGTTTGTAT	7161
DB	1739	TTATGTAGCCTGTGATAAATATGGGCACCTCCATTAGTGTGTTTCAGCAGTCGGGCTCAGGT	1798
QY	7162	GATCAAGCAAAATCCTGTACACAGCCCTATGGGGAGATCTACATGGATACCAACCCCAACTT	7221
DB	1799	CATAAAGAGATTCTGTATACACCTTATGGTGACATCTACCATGATACATACCATCCCTGACTT	1858
QY	7222	TCAGATCATATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACAT	7281
DB	1859	TGAGGTTCATCTGGTTTTCATGGAGGACTCTATGATTTTCTTACTAACTAGTGCATCT	1918
QY	7282	GGGCCGGGAGATTATGATGTGTCTGGCCGACGCTGGACTAGCCAGACCACGAGCTGTG	7341
DB	1919	AGGGCAAAGAGATTATGATGTGTGTCTGGCAGATGGACAACGCCTAATCATCATCATATG	1978
QY	7342	GAAGCACCTTAGTAGCAGCAACGTCATGCCTTTTAACTCTATATGTTTCAAAAACAACAA	7401
DB	1979	GAAACAGTTGA---ACCTTCTTCCCTAAACCATTCACCTCTACTCCTTTTGAGATAACTA	2035
QY	7402	CCCCATCAGCAACTCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCT	7461
DB	2036	CCCAGTTGGCAAAATCAAGATGTTGCCAAAGTATACCAAGATATCGGCACCTGGCTGGA	2095
QY	7462	CACCTTTGGATTCCAGCTACACAACGTGATCCCTGGTTATCCCAACCAAGACATGGATGC	7521
DB	2096	GCTCTTTGGTTTCCAGTTACATAACGTACTACCTGGATTTCCCAACCAAGAACTTGAAAA	2155
QY	7522	CATGGAAACCTCTACGAGTTCATCCACACAGATGAAAAACGCAGGAGTGGGACAAACAG	7581
DB	2156	CATGGAATTAACTTATGAGCTTCTTCAACTGCAGACAAAAACCCAGAGTGGGATCTCTGG	2215
QY	7582	CAAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGTCACTT	7641
DB	2216	AAAGATGATCTCTGGCAATTCAAGTGTGAGCTCCAGAAAAACAATAAGGAATTTCAATTCCTT	2275

QY	7642	AGAACGGTTTGACCAAGCTCTATGGCTCCACAATACACCAGCTGCCAGAGCGTCCAAAGAC	7701
Db	2276	GGATCAGCTTCCTATGACTCCTCAGTACAATAGAGGGAGGTCTCTTGAAGGAGGAAACA	2335
QY	7702	CAAGAAGTTTGCATCCAGCGGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGCTTGAAGGA	7761
Db	2336	GCCGAGGTTTGCTGCTGTCCCTTCTGTCTTTGGCAAGGTATCAAAATTTGCCATCAAGGA	2395
QY	7762	TGGCCGAGTGACCAACAGACATCATCAGTGTGGCCAATGAGGATGGCGAAGGGTTGCTGC	7821
Db	2396	GGGCATAGTGACAGCTGATATTATAGGAGTAGCCAATGAAGATAGCAGCGCTCTTGCTGC	2455
QY	7822	CATCTTGAACCATGCCACTACCTAGAGAACCTGCACCTTCAACATTGATGGGTGGATAC	7881
Db	2456	CATTCTCAACAATGCTCATTTACCTGGAAACCTGCATTTTACCATAGAAGGGAGGGACAC	2515
QY	7882	CCATTACTTTGTGAAACCAGGACCTTCAGAAGGTGACCTGGCCATCCTGGGCCCTCAGTGG	7941
Db	2516	TCACTACTTTAAGCTCGGGTCTCTGAGGAAGACCTGGTGCTCATCGGTAACACTGG	2575
QY	7942	GGGGCGGCAACCTGGAGAAATGGGGTCAACGTCACTGTGCCAGATCAACACACAGTACT	8001
Db	2576	GGGGAGCGGATTCTGGAGAAATGGTGTAAATGTCACTGTGTCTCAGATGACGTCACTGTT	2635
QY	8002	TAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAGTACGGGGCACTGTGCTTGAA	8061
Db	2636	GAATGGGAGGACTAGACGGTTTGCAGATATCCAACCTGCAGCACGGAGCTCTGTCTCAA	2695
QY	8062	CACACGCTACGGGACAACGTTGGATGAGGAGAAGGCACGGGTCTCTGGAGCTGGCCCCGCA	8121
Db	2696	CATCCGGTATGAAACGACTGTCTGAAGAGGAGAAGAACCACTGTGTGGAGATGCCAGACA	2755
QY	8122	GAGAGCCGTGGCCCAAGCGTGGGCCCGCCGACGACAGAGACTGCGGGGAACGGGAGGAAGG	8181
Db	2756	GCGCGCAGTGGCCCCAGGCCCTGGACTCAGGAGCAGAGGAGGCTACAGGAAGGGGAAGAGG	2815
QY	8182	CCTGCGGGCTTGACAGAGGGGAGAGCAGCAGGTGCTGAGCACAGGGCGGTGCAAGG	8241
Db	2816	GACTCGGGTATGGACAGAGGGAGAAAAGCAGCAGCTTTTGGGCACCGGGAGGGTGCAGGG	2875
QY	8242	CTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTTCAGACAGCGCCAA	8301
Db	2876	GTATGATGGGTATTTTGTCTTGTCTGTTGAGCAGTATTTAGAACTTTTCAGACAGTGCCAA	2935
QY	8302	CAACATCCACTTCATGAGACAGAGCGAGATGGGCCCGGAGGTGACAGA	8348
Db	2936	CAATATTCACCTTCATGAGACAGAGTGAATAGGCAGGAGGTAAACAAA	2982

RESULT 8  
AK039472

LOCUS	DEFINITION
1	...
2	...
3	...
4	...
5	...
6	...
7	...
8	...
9	...
10	...
11	...
12	...
13	...
14	...
15	...
16	...
17	...
18	...
19	...
20	...
21	...
22	...
23	...
24	...
25	...
26	...
27	...
28	...
29	...
30	...
31	...
32	...
33	...
34	...
35	...
36	...
37	...
38	...
39	...
40	...
41	...
42	...
43	...
44	...
45	...
46	...
47	...
48	...
49	...
50	...
51	...
52	...
53	...
54	...
55	...
56	...
57	...
58	...
59	...
60	...
61	...
62	...
63	...
64	...
65	...
66	...
67	...
68	...
69	...
70	...
71	...
72	...
73	...
74	...
75	...
76	...
77	...
78	...
79	...
80	...
81	...
82	...
83	...
84	...
85	...
86	...
87	...
88	...
89	...
90	...
91	...
92	...
93	...
94	...
95	...
96	...
97	...
98	...
99	...
100	...

AK039472 1284 bp mRNA linear HTC 19-SEP-2003  
Mus musculus adult male spinal cord CDNA, RIKEN full-length  
enriched library, clone:A330048C04 product:neuregulin 1, full  
insert sequence.

ACCESSION

AK039472.1 GI:26087188

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM	Mus musculus
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

**AUTHORS** Carninci, P. and Hayashizaki, Y.

TITLE	High-efficiency full-length cDNA cloning
-------	--

JOURNAL  
Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED  
10349636

## REFERENCE

REFERENCES  
AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

# TITLE

JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	Query Match	8.9%; Score 743.8; DB 11; Length 1284;
MEDLINE	20499374	Best Local Similarity	86.3%; Pred. No. 1.7e-140;
PUBMED	11042159	Matches	845; Conservative 0; Mismatches 107; Indels 27; Gaps 1;
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE	5		
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE	6 (bases 1 to 1284)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.		
FEATURES	Location/Qualifiers		
source	1. .1284		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="FANTOM DB:A330048C04"		
	/db_xref="MGI:2403942"		
	/db_xref="taxon:10090"		
	/clone="A330048C04"		
	/sex="male"		
	/tissue_type="spinal cord"		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"		
	/dev_stage="adult"		
misc_feature	1. .1284		
	/note="neuregulin 1 (MGD MG1:96083, MGD MG1:1345182, GB NM_011858, evidence: BLASTN, 99%, match=186)"		
ORIGIN			

```
RESULT 9
AL045768
LOCUS
DEFINITION DKFZp434F206.r1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045768
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 718)
JOURNAL Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
COMMENT EST (Koehrer, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
sl sequence also available.
This clone (DKFZp434F206) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source Location/Qualifiers
1..718
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434F206"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"
ORIGIN
Query Match 8.5%; Score 713.8; DB 9; Length 718;
Best Local Similarity 99.7%; Pred. No. 1.6e-134;
Matches 715; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 7431 GCTTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGTACACAAACGTGA 7490
Db 1 GCTTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGTACACAAACGTGA 60
QY 7491 TCCTGGTTATCCCAAACACAGACATGGATGCCATGGAACCTCTCTACGAGTCCATCCACA 7550
Db 61 TCCCTGGTTATCCCAAACACAGACATGGATGCCATGGAACCTCTCTACGAGTCCATCCACA 120
QY 7551 CACAGATGAAAACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAG 7610
Db 121 CACAGATGAAAACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAG 180
QY 7611 TACAGAAGCAGCTCAAGGCCCTTTGTACACCTTAGAACGTTTGACCAGCTCTATGGCTCCA 7670
Db 181 TACAGAAGCAGCTCAAGGCCCTTTGTACACCTTAGAACGTTTGACCAGCTCTATGGCTCCA 240
QY 7671 CAATCACCAGCTGCCAGCAGGTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCT 7730
Db 241 CAATCACCAGCTGCCAGCAGGTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCT 300
QY 7731 TTGGCAAGGGGGTCAAGTTTGCTTGAAGGATGGCCAGTGGCCAGTGACCACAGACATCATCAGTG 7790
Db 301 TTGGCAAGGGGGTCAAGTTTGCTTGAAGGATGGCCAGTGGCCAGTGACCACAGACATCATCAGTG 360
QY 7791 TGGCCAATGAGGATGGGCGAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGA 7850
|||||
```

```
Db 361 TGGCCAATGAGGATGGGCGAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACTAGAGA 420
QY 7851 ACCTGCACCTTCAACCATTTGATGGGGTGGATACCCATTACTTTGTGAAACACGAGACCTTCAG 7910
|||||
Db 421 ACCTGCACCTTCAACCATTTGATGGGGTGGATACCCATTACTTTGTGAAACACGAGACCTTCAG 480
|||||
QY 7911 AAGGTACCTGGCCATCCTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAAATGGGTCA 7970
|||||
Db 481 AAGGTACCTGGCCATCCTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAAATGGGTCA 540
|||||
QY 7971 ACGTCACTGTGTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACA 8030
|||||
Db 541 ACGTCACTGTGTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACA 600
|||||
QY 8031 TCCAGCTCCAGTACGGGGCACTGTCTTGAACACACAGCTACGGGACAAACGTTGGATGAGG 8090
|||||
Db 601 TCCAGCTCCAGTACGGGGCACTGTCTTGAACACACAGCTACGGGACAAACGTTGGATGAGG 660
|||||
QY 8091 AGAAGGCACGGGTCTCTGGAGTGGCCCCGGCAGAGAGCGGTGCGCAAGCGTGGGCC 8147
|||||
Db 661 AGAAGGCACGGGTCTCTGGAGTGTCCCGGCAGAGAGCGGTGCGCAAGCGTGGGCC 717
|||||
RESULT 10
CF534506 856 bp mRNA linear EST 12-SEP-2003
LOCUS UI-M-G10-chg-a-23-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:30536758 5', mRNA sequence.
ACCESSION CF534506
VERSION CF534506.1 GI:34586474
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 856)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1..856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30536758"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Ronald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
```



QY 6744 TACGGTATGACATCCGGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATG 6803  
|||||  
Db 241 TACGGTATGACCTCCGTGACCGCATCACTAGGCTGGGTGATGTACAGTACAAGATGGATG 300  
|||||  
QY 6804 AGGATGGCTTCTGAGGACGGGGGGGTGATATCTTTGAGTACAACCTCAGCTGGCCTGC 6863  
|||||  
Db 301 AGGATGGCTTCTGAGGACGGGGGGGTGGGATGTCTTCGAGTACAATTCAGCCGGCTGC 360  
|||||  
QY 6864 TCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGTGAGGTACCGCTACGATGGCCTGG 6923  
|||||  
Db 361 TCATCAAGGCCTACAACCGGGCTAGTGGTGGAGTGTGAGGTACCGGATACGATGGGCTGG 420  
|||||  
QY 6924 GCGGGCGGTGTCAGCAAGAGCAGCCACAGCCACACCTGCAGTTCCTCTATGCAGACC 6983  
|||||  
Db 421 GACGCCGAGTATCCAGCAAGAGCAGCCACAGCCACACCTGCAGTTCCTTACGCAGACC 480  
|||||  
QY 6984 TGACCAACCCCAAGGTCACCCACCTGTACAAACACTCCAGCTCTGAGATTACATCCC 7043  
|||||  
Db 481 TGACCAACCCCAAGGTCACCCACCTCTACAACCACTCCAGCTCTGAGATTACATCCC 540  
|||||  
QY 7044 TCTACTACGACTTGCAAGGACACCTCTTTGCCATGAGCTGAGCAGTGGTGATGATTTT 7103  
|||||  
Db 541 TCTACTATGATCTGCAAGGACACCTCTTTGCCATGAGCTGAGCAGTGGGATGATTTT 600  
|||||  
QY 7104 ACATAGCTTGTGACAAACATCGGACCCCTCTTCTGCTCTTTAGTGAACAGTTTGATGA 7163  
|||||  
Db 601 ACATAGCTTGTGACAAACATCGGACCCCGCTTCTGCTCTTCAGTGAACCTGGCTTGATGA 660  
|||||  
QY 7164 TCAAGCAATCCTGTACACAGCCTATGGGGAGATCTACATGGATACCAACCCCACTTTC 7223  
|||||  
Db 661 TTAAGCAGATCCTATACACGGGCTATGGGAGATCTACATGGACA-CANACCCANCTTCC 719  
|||||  
QY 7224 AGATCATATAGGCTACCATGTTGGCCTCTATGATCCACTCACCAGCTTGTCCACATGG 7283  
|||||  
Db 720 AGATCATATCGGCTA-CACGGCGNCTCTATGAT---CACTCACANGCTTGTACATGG 775  
|||||  
QY 7284 GCCGGCGAGATTATGATGTGCTGGCCGACGCTGGACTAGCC 7325  
|||||  
Db 776 GCCACGGGA-TATGATGTGCTGGCTGG-CGCTGGACAAGCC 815  
|||||

RESULT 12  
BG036207 870 bp mRNA linear EST 24-JAN-2001  
LOCUS 602326960F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4428351 5',  
DEFINITION mRNA sequence.  
ACCESSION BG036207  
VERSION BG036207.1 GI:12431132  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 870)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10178 row: h column: 16  
High quality sequence stop: 713.  
FEATURES  
source location/Qualifiers  
1..870  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clone="IMAGE:4428351"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_91"  
/note="Organ: prostate; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.4 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 7.5%; Score 629; DB 10; Length 870;  
Best Local Similarity 98.2%; Pred. No. 3.7e-117;  
Matches 700; Conservative 0; Mismatches 5; Indels 8; Gaps 6;  
QY 904 AGGCACCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACTGACGTCCAGCAGT 963  
|||||  
Db 1 AGGCACCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACTGACGTCCAGCAGT 59  
|||||  
QY 964 GTACTCTCTCCGCCCCGACCCCTGCCCGCAGCACCTTGCCTCCGCGCCCTTTAACTT 1023  
|||||  
Db 60 GTACTCTCTCCGCCCCGACCCCTGCCCGCAGCACCTTGCCTCCGCGCCCTTTAACTT 119  
|||||  
QY 1024 CAAGAAAGCCCTCCAAGTACTGTAACTGGAAGTGGCGACCCCTGAGCGCCATCGTCATCTC 1083  
|||||  
Db 120 CAAGAAAGCCCTCCAAGTACTGTAACTGGAAGTGGCGACCCCTGAGCGCCATCGTCATCTC 179  
|||||  
QY 1084 AGCACTCTGGTCACTCCTGTGGCATACTTTGTGGCCATGCACCTGTTGGCCTAACTG 1143  
|||||  
Db 180 AGCACTCTGGTCACTCCTGTGGCATACTTTGTGGCCATGCACCTGTTGGCCTAACTG 239  
|||||  
QY 1144 GCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGAGGACACAGCCAGCAGTTG 1203  
|||||  
Db 240 GCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGAGGACACAGCCAGCAGTTG 299  
|||||  
QY 1204 GCCTGTGCCAACCGACGCTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCCCTGA 1263  
|||||  
Db 300 GCCTGTGCCAACCGACGCTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCCCTGA 359  
|||||  
QY 1264 CAGGAAAGGCAAGGAACCCACAGAAGGAAGCCAGTAGTTCTTTCCAGAGGACAGTTT 1323  
|||||  
Db 360 CAGGAAAGGCAAGGAACCCACAGAAGGAAGCCAGTAGTTCTTTCCAGAGGACAGTTT 419  
|||||  
QY 1324 CATAGATTCTGGAGAAATTGATGTGGGAAGGCGAGCCTCCAGAAAGATTCTCTCTGGCAC 1383  
|||||  
Db 420 CATAGATTCTGGAGAAATTGATGTGGGAAGGCGAGCCTCCAGAAAGATTCTCTCTGGCAC 479  
|||||  
QY 1384 TTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCTCT 1443  
|||||  
Db 480 TTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCTCT 539  
|||||  
QY 1444 GGGAAAGCAGCCCT--GGTTGGCATTTATGGCAGAAAAGSGCCTCCCTTCCACA-TAC 1500  
|||||  
Db 540 GGGAAAGCAGCCCTGGGGTTGGCATTTATGGCAGAAAAGSGCCTCCCTTCCACATTAC 599  
|||||  
QY 1501 ACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCT-CCTAACCCAGGAGGCGCGGA 1559  
|||||  
Db 600 ACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCCTAACCCAGGAGGCGCGGA 659  
|||||  
QY 1560 GCCTAGA-GGGGACCCCGGCCAGTCT--CGGGGAAGTGTGCCCCCTCCAGC 1609  
|||||  
Db 660 GCCTAGAGGGGACCCCGGCCAGTCTTCGGGGGAAGTGTGCCCCCTTCCCGC 712  
|||||

RESULT 13  
CF536185 728 bp mRNA linear EST 12-SEP-2003  
LOCUS UI-M-GIO-chh-a-16-0-UI.r1 NIH\_BMAP\_G10 Mus musculus cDNA clone  
DEFINITION IMAGE:30534831 5', mRNA sequence.  
ACCESSION CF536185  
VERSION CF536185.1 GI:34588165  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 728)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousefl.html>  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
Seq primer: pYX-5.  
Location/Qualifiers  
1. 728  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30534831"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP GiO"  
/note="Organ: Brain; Vector: pYX- Asc; Site\_1: EcoR I;  
Site\_2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is AGCAGACAG. This library was created for the University  
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

ORIGIN

Query Match 7.3%; Score 613.6; DB 14; Length 728;  
Best Local Similarity 90.3%; Pred. No. 4.5e-114;  
Matches 655; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 2762 ATCCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTATTCGTGGCCAAAGTGATG 2821  
Db 4 ATCCCCGGGGAGAACCCCTTTGACGGAGGGCATGCTTGTGTATTCGTGGAGAGTGATG 63  
QY 2822 ACATCAGATGGAAACCCCTCGTTGGTGTGAACATCATGTTTGTCAATAACCCCTCTTTT 2881  
Db 64 ACGTCAGATGGAGCCCCCATTTGGTTGGCTGAACATCATGTTTTCATCAATAACCCCTCTTTT 123  
QY 2882 GGATATACATCAGCAGGCAAGATGGCAGCTTTTGACTTGGTGACAAATGGCGGATCTCC 2941  
Db 124 GGATACACAATCAGCAGGCAAGATGGCAGCTTTTGACCTGGTCACAAATGGCGGATCTCC 183  
QY 2942 ATCATCTCGGTTGTCAGCGGGCACCTTTTCATCACACAGGAGCACACCCTGTGGTGCCA 3001  
Db 184 ATCATCTTGAGATTCGAAAGGGCACCTTTTCATCACACAGGAGCATACCTCTGTGGCTACCC 243  
QY 3002 TGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGAGAGATGAGATCCC 3061  
Db 244 TGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGAGAGATGAGATCCC 303  
QY 3062 AGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTCGTCCTCTCCATCCCCACTGACG 3121  
Db 304 AGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTCGTCCTCTCCATCCCCACTGACA 363

QY 3122 TCCTTCGCCAGCTCCTGTGTCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCTTTGCAG 3181  
Db 364 TCCTTCGCCAGCTCCTGTGTCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCTTTGCAG 423  
QY 3182 GAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGTACCTGAGCAGCCGACCCCT 3241  
Db 424 GAGGAAATCGTCATCGCTGGCTGCAAGATGAGGCTGAGTACCTGAGCAGCCGACCTCT 483  
QY 3242 GGCTACAAATCTGCTCTGAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATG 3301  
Db 484 GGCTATAAGTCTGCTCTGAGGATCAGTCTCACACCCACCCACCATCCCTTCAACCTCATG 543  
QY 3302 AAGGTGCACCTCATGCTAGCGGTGAGGGCGCCCTCTTTCAGGAAGTGGTTCGCTGCAGCC 3361  
Db 544 AAGGTCCACCTCATGCTAGCGGTGAGGGCGCCCTCTTTCAGGAAGTGGTTCGCGCAGCT 603  
QY 3362 CCAGACCTGCTCTATTATTTCATTTGGGACAAAGACAGACGCTCTACAACCAAGGTGTTT 3421  
Db 604 CCTGACCTGCTCTATTATTTCATTTGGGACAAAGACAGATGTCTACAACCAAGGTGTTT 663  
QY 3422 GGGCTTTTCAGAAAGCCTTTGTTTCCGTGGTTATGAATATGAATCCTGCCAGATCTAATC 3481  
Db 564 GGACTCTCAGAGGCTTTGTTTCCGTGGTTATGAGTATGATCCTGCCAGATCTGATC 723  
QY 3482 CTGTG 3486  
Db 724 CTGTG 728

RESULT 14

CB524534  
LOCUS  
DEFINITION  
UI-M-FY0-cey-b-20-0-UI.r1 NIH\_BMAP\_FY0 Mus musculus cDNA clone  
IMAGE:6843669 5', mRNA sequence.  
ACCESSION  
CB524534  
VERSION  
CB524534.1 GI:29357889  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 770)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousefl.html>  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. 770  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6843669"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_FY0"  
/note="Organ: Brain; Vector: pYX- Asc; Site\_1: EcoR I;  
Site\_2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT

FEATURES

source



Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES  
source  
1. .2627  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:D030019H23"  
/db\_xref="MGI:2418637"  
/db\_xref="taxon:10090"  
/clone="D030019H23"  
/tissue\_type="whole body"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="9 days embryo"  
misc\_feature  
1. .2627  
/notes="odd Oz/ten-m homolog 3 (Drosophila)  
(MGI|MGI:1345183, GB|NM\_011857, evidence: BLASTN, 99%,  
match=1753)"

ORIGIN		Query Match		7.2%; Score 602.6; DB 11; Length 2627;		Best Local Similarity 64.3%; Pred. No. 1.8e-111;		Matches 980; Conservative 0; Mismatches 494; Indels 51; Gaps 3;	
QY	880	CGGGCACTTCCCTTCAAGCCTGG	---	AGGCACCTCCCGCTCTTCTGCACCATCACC	936				
Db	317	CAGGCATTTCCCTATTCAAACACAGGACAGGACGACGCCACTGTTCAGTACGGCAACCCC	376						
QY	937	AGGGTACCCACTGACGTCCAGCACAGTGTACTCTCCTCCGCCCGACCCCTGCCCCGAG	996						
Db	377	GGGATACACAAATGGCATCTGGCTCTGTTTATTTCTCCGCTACCCGCCACTTCCTAGAAA	436						
QY	997	CACCTTCGCCCGCGCGCCTTTAAACCTCAAGAAAGCCCTCCAAAGTGTAACTGTAAGTGAAGTG	1056						
Db	437	CACCCATATCAAGAAAGTGTCTTTAAATTCAGAAAGTCTTCAAAGTACTGCAGCTGGAGGTG	496						
QY	1057	QGCAGCCCTGAGCGCCATCGTTCATCTCAGCCACTCTGGTTCATCCTGCTGGCATACTTTGT	1116						
Db	497	CACCGCACTGTGTGTGTAGGGGTCTCAGTGTCTCTGGCCATTCCTCTCTCTATTTAT	556						
QY	1117	GGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGAGGGGCGAGATGTATGA	1176						
Db	557	AGCAATGCATCTATTGTGGCTCAACTGGCACTTACAGCAGACGCGAAATGACACATTGCA	616						
QY	1177	GATCAGGAGGACACAGCCAGCAGTGTGGCTGTGCCAACCGACGTCTCCCTATACCCCTC	1236						
Db	617	GA-----ATGGAAGTGAATTCCTGACACCGTGCACAAACAACTGTATCGTTACCTTC	670						
QY	1237	AGGGGGCACTGGCTTAGAGACCCCTGCAGGAAAGGCAAGGAACACAGAAAGAGCC	1296						
Db	671	TGGCGACA-----ATGGAATAAT	688						
QY	1297	CAGTAGTTTCTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGGAGGCG	1356						
Db	689	AGGTGGATTACACATGAAATAACACCATAGATTCCGGAGAACTTGATATTGGCCGGAG	748						
QY	1357	AGCCTCCAGAAATTCCTCCTCGCACTTCTGGAGATCTCAAGTGTTCATAGACCATCC	1416						
Db	749	AGCAATTCAAGAGGTTCCCCCGGGATCTCTGGAGATCGCAGCTCTTTATTGATCAGCC	808						
QY	1417	TGTGCATCTGAATTCATGTGTCTCTGGGAAAGGACGCCCTGGTGGCATTATGGCAG	1476						
Db	809	ACAGTTTCTTAAGTTCAACATCTCTCTTCAGAAAGGATGCATTGATCGGAGTGTACGGCCG	868						
QY	1477	AAAGGCCTCCCTCCTTCACATACACAGTTTGTGAGCTGTGGATGGCAGGAG	1536						
Db	869	GAAGGGCTTACCGCCTTCCCATACTCAGTACGACTTTGTGGAACTACTGGATGGTAGCAG	928						
QY	1537	GCTCCTAACCCAGAGCGCGGAGCCTTAGAGGGGACCCCGCCAGTCTCGGGAACTGT	1596						
Db	929	GTTAATTGCGAGAGCAGCGGAACCTGGTGGAGTCCGAAAGACCGCGCGGCGGAGCGAG	988						

QY	1597	GCCCCCTCCACCCATGAGACAGGCTTCCATCCAGTATTTGGATTCCAGGAATCTGGCACTT	1656						
Db	989	ATCTGTAGCCTGCACGAAGCTGGCTTCATCCAGTACTTGGATTCTGGAATCTGGCATCT	1048						
QY	1657	GGCTTTTACAAATGACGGAAAGGAGTCAGAGTGGTTTCTTTTCTCACCACTGCCATTGA	1716						
Db	1049	GGCTTTTATAACGACGGGAAAAAACCCAGAGCAGGTCTCTTTAAACAGATCGTTATAGA	1108						
QY	1717	GTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGAATGCTCTCTGGGACCTG	1776						
Db	1109	GTCGTGGTGAATGCCCCCGAAATGCCATGGAAATGGAGAGTGTGTTTCTGGAACCTG	1168						
QY	1777	CAACTGCTTCTGCGTTTCTGGGCCCCGACTGTGGCAGAGCCTCCTGCCCGTGTCTG	1836						
Db	1169	CCATTGTTTCCCGGGTTTCTAGGTCGGATTGTTCAAGAGCAGCCTGTCCGGTGTCTG	1228						
QY	1837	TAGCGGAATGCGCAATACATGAAAGGCAGATGCTTGTGCCACAGTGGCTGGAAAGGCGC	1896						
Db	1229	TAGTGGCAACGGGCAATACTCCAAAGGGCGCTGCCCTGTCTTCAAGTGGTGGAGGGCAC	1288						
QY	1897	TGAGTGCATGTGCCCCACCAACCAAGTGTATCGATGTGGCCTGCAGCAACCATGGCACCTG	1956						
Db	1289	CGAGTGTGACGTGCCGACGACCCAGTGCATTCACCCGAGTGGGGGTCTGTGGGATTG	1348						
QY	1957	CATCAGGGGCACTGCATCTGCAACCCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGA	2016						
Db	1349	CATCATGGGCTCTTGGCTTGTAACTCGGGATACAAAGGAGAAAACCTGTGAGGAAGCGGA	1408						
QY	2017	CTGCATGGACCCACATGTTTCAGGCGGGGTGTCTGCTGAGAGGCGAATGCCATTGCTT	2076						
Db	1409	CTGTCTAGACCTGGATGTTCTAATCACGGGGTGTGTATCCATGGGGAATGTCACTGCAA	1468						
QY	2077	TGTGGATGGGAGGACCAACTGCGAGACCCCGAGGCCACATGCTTAGACCAGTGTTC	2136						
Db	1469	TCCAGGCTGGGTGGCAGCAACTGTGAAATACTGAAGACTATGTGTGCAGACCAAGTGTCTC	1528						
QY	2137	AGGCCACGGAACTTCTCCCGGACACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGG	2196						
Db	1529	AGGCCACGGGACTTACCTTCAAGAAAGCGGCTCCTGCACCTTGGCACCCCAATTTGACTGG	1588						
QY	2197	ACACGACTGTTCTATCGAGATCTGTGTGCTCCGACTGTGTGGCCATGGCGTGTGCGTAGG	2256						
Db	1589	CCCGACTGCTCAAATGAAATATGTTCACTGGACTGCGGCTCACACGGCGTCTGCATGGG	1648						
QY	2257	GGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGGCGAGCCTTGGACCGCGGGCTGCCA	2316						
Db	1649	GGGCTCCTGCTGTGAAGAAAGGCTGGACCGGCCCGCGCTGTAATCAGAGAGCTTGCCA	1708						
QY	2317	CCCGCGCTGTGCCGAGCATGGGACCTGCCCGGACGGCAAGTGGAGTGCAGCCCTGGCTG	2376						
Db	1709	CCCTCGCTGTGTGAGCACGGGACGTTGCAAGGACGGCAAGTGGAGTGCAGCCCAAGGATG	1768						
QY	2377	GAATGGCGAACACTGCACCATCGCT	2401						
Db	1769	GAACGGAGAGCACTGCACAAATTGGT	1793						

Search completed: June 25, 2004, 02:14:52  
Job time : 18822 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 17:03:54 ; Search time 2140 Seconds  
(without alignments)  
17883.374 Million cell updates/sec

Title: US-10-029-020-13  
Perfect score: 8354  
Sequence: 1 gtttgtggtgaggagc.....ccggaggtgacagaggac 8354

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 2290544650 residues

Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA: \*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq: \*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq: \*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq: \*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq: \*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq: \*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq: \*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq: \*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq: \*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq: \*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq: \*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq: \*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq: \*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2: \*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq: \*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq: \*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq: \*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq: \*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq: \*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8354	100.0	8354	13	US-10-383-201-43
2	8354	100.0	8354	13	US-10-029-020-13
3	7991.2	95.7	8355	13	US-10-383-201-55
4	7779.2	93.1	8438	13	US-10-042-865-1
5	3502.2	41.9	8645	17	US-10-038-854-37
6	3473.8	41.6	8675	17	US-10-038-854-35
7	3422	41.0	8473	17	US-10-038-854-39
8	3398.8	40.7	8487	17	US-10-038-854-41
9	3397.8	40.7	8689	9	US-09-808-602-78
10	3397.8	40.7	8689	10	US-09-800-198-66
11	3385	40.5	9826	9	US-09-808-602-7
12	3385	40.5	9826	10	US-09-800-198-7
13	3384.8	40.5	8797	9	US-09-808-602-74
14	3384.8	40.5	8797	9	US-09-808-602-77

15	3384.8	40.5	8797	10	US-09-800-198-62	Sequence 62, Appl
16	3384.8	40.5	8797	10	US-09-800-198-65	Sequence 65, Appl
17	3351.4	40.1	9729	9	US-09-808-602-12	Sequence 12, Appl
18	3351.4	40.1	9729	10	US-09-800-198-12	Sequence 12, Appl
19	3258	39.0	12880	16	US-10-295-027-927	Sequence 927, App
20	3165	37.9	9695	16	US-10-144-194A-81	Sequence 81, Appl
21	3158.4	37.8	8575	13	US-10-072-012-143	Sequence 143, Appl
22	3070.8	36.8	8409	9	US-09-808-602-79	Sequence 79, Appl
23	3070.8	36.8	8409	10	US-09-800-198-67	Sequence 67, Appl
24	3068.6	36.7	9058	16	US-10-144-194A-79	Sequence 79, Appl
25	2533	30.3	6560	9	US-09-808-602-76	Sequence 76, Appl
26	2533	30.3	6560	10	US-09-800-198-64	Sequence 64, Appl
27	1708.8	20.5	3111	9	US-09-773-517-12	Sequence 12, Appl
28	1708.8	20.5	3111	9	US-09-792-025-12	Sequence 12, Appl
29	1708.8	20.5	3111	9	US-09-849-868-12	Sequence 12, Appl
30	1708.8	20.5	3111	15	US-10-290-578-1	Sequence 1, Appl
31	1708.8	20.5	3111	15	US-10-453-183-12	Sequence 12, Appl
32	1673.6	20.0	1680	15	US-10-290-578-3	Sequence 3, Appl
33	1560.4	18.7	3614	13	US-10-342-887-1743	Sequence 1743, Ap
34	1560.4	18.7	3614	13	US-10-172-118-1743	Sequence 1743, Ap
35	1378.2	16.5	1399	13	US-10-383-201-49	Sequence 49, Appl
36	1378.2	16.5	1399	13	US-10-383-201-59	Sequence 59, Appl
37	1377.2	16.5	1476	13	US-10-383-201-41	Sequence 41, Appl
38	1375.2	16.5	1392	13	US-10-383-201-45	Sequence 45, Appl
39	1375.2	16.5	1392	13	US-10-383-201-53	Sequence 53, Appl
40	1309.4	15.7	1371	13	US-10-383-201-51	Sequence 51, Appl
41	1185.6	14.2	2387	15	US-10-290-578-11	Sequence 11, Appl
42	833	10.0	3217	15	US-10-198-846-13976	Sequence 13976, A
43	807.2	9.7	829	13	US-10-383-201-47	Sequence 47, Appl
44	807.2	9.7	829	13	US-10-383-201-61	Sequence 61, Appl
45	744	8.9	802	13	US-10-383-201-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1  
US-10-383-201-43  
; Sequence 43, Application US/10383201  
; Publication No. US20040029226A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook II, John et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHC  
; FILE REFERENCE: 21402-568A  
; CURRENT APPLICATION NUMBER: US/10/383,201  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: 10/029020  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/365,984  
; PRIOR FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/372,022  
; PRIOR FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 60/389,143  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/391,779  
; PRIOR FILING DATE: 2002-06-26  
; PRIOR APPLICATION NUMBER: 60/410,755  
; PRIOR FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 60/412,957  
; PRIOR FILING DATE: 2002-09-23  
; PRIOR APPLICATION NUMBER: 10/051,874  
; PRIOR FILING DATE: 2002-01-16  
; PRIOR APPLICATION NUMBER: 60/366,928  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 10/055,877  
; PRIOR FILING DATE: 2002-01-22  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 43  
; LENGTH: 8354  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: CDS									
; LOCATION: (35) .. (8341)									
US-10-383-201-43									
Query Match									
Best Local Similarity 100.0%; Score 8354; DB 13; Length 8354;									
Matches 8354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GTTTGTGATCTGGAGGAGCGCGCGCGGAGGCGCCATGGACGTGAAGGAGAGGAGCCCTTA	60						
Db	1	GTTTGTGATCTGGAGGAGCGCGCGCGGAGGCGCCATGGACGTGAAGGAGAGGAGCCCTTA	60						
QY	61	CCGCTCGCTGACCCCGCGCGCGGAGCGCGGCTACACCAAGCTCGTCCCGCGGACAG	120						
Db	61	CCGCTCGCTGACCCCGCGCGCGGAGCGCGGCTACACCAAGCTCGTCCCGCGGACAG	120						
QY	121	CGAGGAGGCGCAAGCCCGCGCAGAAATCGTACAGCTCCAGCGAGACCCTGAAGGCTACGA	180						
Db	121	CGAGGAGGCGCAAGCCCGCGCAGAAATCGTACAGCTCCAGCGAGACCCTGAAGGCTACGA	180						
QY	181	CCAGGACGCGCGCTAGCCTATGGCAGCGCGCTCAAGGACATTTGTCCCGAGGCGCGA	240						
Db	181	CCAGGACGCGCGCTAGCCTATGGCAGCGCGCTCAAGGACATTTGTCCCGAGGCGCGA	240						
QY	241	GGAATTCTGCGGCACAGGTGCCAATTCAACCTGCGGGAGTGGGGCTGGAAGAAGTAAC	300						
Db	241	GGAATTCTGCGGCACAGGTGCCAATTCAACCTGCGGGAGTGGGGCTGGAAGAAGTAAC	300						
QY	301	GCCCCCTCACGGGACCCTGTACCGGACAGACATTTGGCCTGCCCCCAATGCGGCTACTCCAT	360						
Db	301	GCCCCCTCACGGGACCCTGTACCGGACAGACATTTGGCCTGCCCCCAATGCGGCTACTCCAT	360						
QY	361	GGGGCTGGCTGTATGCCGACATGGAGGCTGACACGGTGTCTCCCTGAGCACCCCGT	420						
Db	361	GGGGCTGGCTGTATGCCGACATGGAGGCTGACACGGTGTCTCCCTGAGCACCCCGT	420						
QY	421	CGCTCTGTGGGCGGAGCACACGCTCAGGGCGCAGCTCCTGCTGTCCAGCCGGSCAA	480						
Db	421	CGCTCTGTGGGCGGAGCACACGCTCAGGGCGCAGCTCCTGCTGTCCAGCCGGSCAA	480						
QY	481	TTCCAATCTCACACTCACCGACACCGAGCATGAAACACTGAGACTGATCATCCGGGCGG	540						
Db	481	TTCCAATCTCACACTCACCGACACCGAGCATGAAACACTGAGACTGATCATCCGGGCGG	540						
QY	541	CCTGCAGAACACGCGGGCTCCGACCGCGCGCGCTCTCGACCGCCACACCC	600						
Db	541	CCTGCAGAACACGCGGGCTCCGACCGCGCGCGCTCTCGACCGCCACACCC	600						
QY	601	CAACCGACACACGCGGGCTCCATTAACCTCCCTGAAACCGGGGCAACTTCACGCCGAGGAG	660						
Db	601	CAACCGACACACGCGGGCTCCATTAACCTCCCTGAAACCGGGGCAACTTCACGCCGAGGAG	660						
QY	661	CAACCCGAGCCCGCGCGCCACCGACCACTCGCTCTCCGGAGAGCCCCCTGCGCGGCGCGC	720						
Db	661	CAACCCGAGCCCGCGCGCCACCGACCACTCGCTCTCCGGAGAGCCCCCTGCGCGGCGCGC	720						
QY	721	CCAGGAGCCTGCCACCGCCAGGAGAACTGGCTGTCAACAGCAACATCCCCCTGGAGAC	780						
Db	721	CCAGGAGCCTGCCACCGCCAGGAGAACTGGCTGTCAACAGCAACATCCCCCTGGAGAC	780						
QY	781	CAGGAACCTAGGCAAGCAGCCATTCTAGGGACATTTGACGACAACTCATTTGAGATGGA	840						
Db	781	CAGGAACCTAGGCAAGCAGCCATTCTAGGGACATTTGACGACAACTCATTTGAGATGGA	840						
QY	841	CATTCTCGGCGCTCCCGCCATGATGGGCTTACAGTGACGGGCACTTCTCTTCAAGCC	900						
Db	841	CATTCTCGGCGCTCCCGCCATGATGGGCTTACAGTGACGGGCACTTCTCTTCAAGCC	900						
QY	901	TGAGGACACCTCCCGCTCTTCTGCACACATCACAGGGTACCCACTGACGTCCAGCAC	960						
Db	901	TGAGGACACCTCCCGCTCTTCTGCACACATCACAGGGTACCCACTGACGTCCAGCAC	960						
QY	961	AGTGACTCTCTCCGCGCCCGACCCCTGCGCGCGAGCACCTTCCGCGCGCGCTTTAA	1020						

Db	961	AGTGACTCTCTCCGCGCCCGACCCCTGCCCGCAGACCTTCGCCCGCGCGCTTAA	1020						
QY	1021	CCTCAAGAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCCATCGTCAT	1080						
Db	1021	CCTCAAGAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCCATCGTCAT	1080						
QY	1081	CTCAGCCACTCTGGTCTATCCTGCTGGCATACTTTGTGGCCATGCACCTGTTTGGCCTAAA	1140						
Db	1081	CTCAGCCACTCTGGTCTATCCTGCTGGCATACTTTGTGGCCATGCACCTGTTTGGCCTAAA	1140						
QY	1141	CTGGCACCTGCAGCCGATGGAGGGCAGATGTATGATCACGGAGGACACAGCCAGCAG	1200						
Db	1141	CTGGCACCTGCAGCCGATGGAGGGCAGATGTATGATCACGGAGGACACAGCCAGCAG	1200						
QY	1201	TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCCC	1260						
Db	1201	TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCCC	1260						
QY	1261	TGACAGGAAAGGCAAGGAACACACAGAAAGGAAAGCCAGTAGTTTCTTCCAGAGGACAG	1320						
Db	1261	TGACAGGAAAGGCAAGGAACACACAGAAAGGAAAGCCAGTAGTTTCTTCCAGAGGACAG	1320						
QY	1321	TTTCATAGATTCTGGAGAAATTGATGTGGGAAGGGCAGCCTCCAGAAAGATTCTCTCTGG	1380						
Db	1321	TTTCATAGATTCTGGAGAAATTGATGTGGGAAGGGCAGCCTCCAGAAAGATTCTCTCTGG	1380						
QY	1381	CACTTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCATCTGAAATTCAATGTGTC	1440						
Db	1381	CACTTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCATCTGAAATTCAATGTGTC	1440						
QY	1441	TCTGGGAAAGGCGAGCCCTGGTTGGCATTTATGGCAGAAAGGCTCCCTCTTTCACATAC	1500						
Db	1441	TCTGGGAAAGGCGAGCCCTGGTTGGCATTTATGGCAGAAAGGCTCCCTCTTTCACATAC	1500						
QY	1501	ACAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGCTCTTAACCCAGGAGCGCGGAG	1560						
Db	1501	ACAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGCTCTTAACCCAGGAGCGCGGAG	1560						
QY	1561	CCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGTGCCCCCTCCAGCCATGAGACAGG	1620						
Db	1561	CCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGTGCCCCCTCCAGCCATGAGACAGG	1620						
QY	1621	CTTCATCCAGTATTGGATTCAAGAACTTGGCACTTGGCTTTTACAATGACGGAAGGA	1680						
Db	1621	CTTCATCCAGTATTGGATTCAAGAACTTGGCACTTGGCTTTTACAATGACGGAAGGA	1680						
QY	1681	GTCAGAAAGTGGTTTCTTTCTACCACTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA	1740						
Db	1681	GTCAGAAAGTGGTTTCTTTCTACCACTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA	1740						
QY	1741	CTGCTATGGCAATGGTACTGCTCTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGG	1800						
Db	1741	CTGCTATGGCAATGGTACTGCTCTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGG	1800						
QY	1801	CCCCGACTGTGGCAGAGCCTCTGCCCGCTGTCTGTAGCGGAAATGGCCAAATACATGAA	1860						
Db	1801	CCCCGACTGTGGCAGAGCCTCTGCCCGCTGTCTGTAGCGGAAATGGCCAAATACATGAA	1860						
QY	1861	AGGCAGATGCTTGTGCCACAGTGGCTGGAAAGGCGCTGATGCGGACCTGTGCAACCA	1920						
Db	1861	AGGCAGATGCTTGTGCCACAGTGGCTGGAAAGGCGCTGATGCGGACCTGTGCAACCA	1920						
QY	1921	GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCACGGGCACCTGTGCAAT	1980						
Db	1921	GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCACGGGCACCTGTGCAAT	1980						
QY	1981	CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCTGAGACCCACATGTTGAGG	2040						
Db	1981	CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCTGAGACCCACATGTTGAGG	2040						
QY	2041	CCGGGGTGTCTGCGTGAGAGGCGAATCCATTGCTTTGTGGGATGGGAGGCGACCAACTG	2100						
Db	2041	CCGGGGTGTCTGCGTGAGAGGCGAATCCATTGCTTTGTGGGATGGGAGGCGACCAACTG	2100						

Db 2041 CCGGGTGTCTCGGTGAGAGCGGAATGCCATTGCTTTGTGGATGGGAGGCACCAACTG 2100

Qy 2101 CGAGACCCCCAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAACCTTCTCTCCCGGA 2160

Db 2101 CGAGACCCCCAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAACCTTCTCTCCCGGA 2160

Qy 2161 CACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG 2220

Db 2161 CACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG 2220

Qy 2221 TGCTGCCGACTGTGTGGCCCATGGCGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGG 2280

Db 2221 TGCTGCCGACTGTGTGGCCCATGGCGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGG 2280

Qy 2281 CTGGATGGGGGACGCTGCCAGCCAGCGGGCCCTGCCACCCGCGCTGTGCCGAGCATGGGAC 2340

Db 2281 CTGGATGGGGGACGCTGCCAGCCAGCGGGCCCTGCCACCCGCGCTGTGCCGAGCATGGGAC 2340

Qy 2341 CTGCCGCGACGGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGC 2400

Db 2341 CTGCCGCGACGGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGC 2400

Qy 2401 TCACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGCTGGTGTGCAATGGCAACGGCAG 2460

Db 2401 TCACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGCTGGTGTGCAATGGCAACGGCAG 2460

Qy 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGG 2520

Db 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGG 2520

Qy 2521 CTGTGACACTTCCATGGAGACTGCCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCT 2580

Db 2521 CTGTGACACTTCCATGGAGACTGCCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCT 2580

Qy 2581 GGTGGACTGCATGGACCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640

Db 2581 GGTGGACTGCATGGACCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640

Qy 2641 CCTTGGCTCCCCCTAACCCTCTGGACATCAATCCAGGAGACACAGGTCCTCTGTTCACAGCA 2700

Db 2641 CCTTGGCTCCCCCTAACCCTCTGGACATCAATCCAGGAGACACAGGTCCTCTGTTCACAGCA 2700

Qy 2701 GAACTTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGCGAGGACAGCACGCACAT 2760

Db 2701 GAACTTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGCGAGGACAGCACGCACAT 2760

Qy 2761 AATCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGCCAAAGTGAT 2820

Db 2761 AATCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGCCAAAGTGAT 2820

Qy 2821 GACATCAGATGGAACCCCTTGTGTGTAACATCAGTTTTGTCAATAAACCTCTCTT 2880

Db 2821 GACATCAGATGGAACCCCTTGTGTGTAACATCAGTTTTGTCAATAAACCTCTCTT 2880

Qy 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC 2940

Db 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC 2940

Qy 2941 CATCATCCTGCGGTTCCGAGCGGCACCTTTCATCACACAGGAGCACACCCCTGTGGCTGCC 3000

Db 2941 CATCATCCTGCGGTTCCGAGCGGCACCTTTCATCACACAGGAGCACACCCCTGTGGCTGCC 3000

Qy 3001 ATGGGATCGCTTCTTTGTATGGAACCATCATCATGAGACATGAGGAGATGAGATTCC 3060

Db 3001 ATGGGATCGCTTCTTTGTATGGAACCATCATCATGAGACATGAGGAGATGAGATTCC 3060

Qy 3061 CAGCTGTGACCTGAGCAATTTGCCCGCCCCAACCCAGTCGTCTCTCCATCCCCACTGAC 3120

Db 3061 CAGCTGTGACCTGAGCAATTTGCCCGCCCCAACCCAGTCGTCTCTCCATCCCCACTGAC 3120

Qy 3121 GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCCATTTGTCGGAAATTCAGGCTTTGCA 3180

Db 3121 GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCCATTTGTCGGAAATTCAGGCTTTGCA 3180

Qy 3181 GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC 3240

Db 3181 GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC 3240

Qy 3241 TGGGTACAAATCTGTCTCTAGGATCAGCCTCACCCACCGACCATCCCCTTCAACCTCAT 3300

Db 3241 TGGGTACAAATCTGTCTCTAGGATCAGCCTCACCCACCGACCATCCCCTTCAACCTCAT 3300

Qy 3301 GAAGGTGCACCTCATGTGTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGC 3360

Db 3301 GAAGGTGCACCTCATGTGTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGC 3360

Qy 3361 CCAGACCTGTCTCTATTTATTTTTCATTTGGGACAAGACAGACGCTCTACAACCAGAGGTGT 3420

Db 3361 CCAGACCTGTCTCTATTTATTTTTCATTTGGGACAAGACAGACGCTCTACAACCAGAGGTGT 3420

Qy 3421 TGGCTTTTCAGAAAGCCTTTGTTTCCGTGGGTTATGAAATATGAATCCTGCCAGATCTAAT 3480

Db 3421 TGGCTTTTCAGAAAGCCTTTGTTTCCGTGGGTTATGAAATATGAATCCTGCCAGATCTAAT 3480

Qy 3481 CCTGTGGAAAAAAGAACAAACAGTGTGCAGGGCTATGAAATTCACGCTCCAAAGCTTGG 3540

Db 3481 CCTGTGGAAAAAAGAACAAACAGTGTGCAGGGCTATGAAATTCACGCTCCAAAGCTTGG 3540

Qy 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCCTGCACAAAGG 3600

Db 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCCTGCACAAAGG 3600

Qy 3601 GAATGGGGAGAACCCAGTTTGTCTCAGCAGCCTCCTGTCTATTTGGAGCATCATGGGCAA 3660

Db 3601 GAATGGGGAGAACCCAGTTTGTCTCAGCAGCCTCCTGTCTATTTGGAGCATCATGGGCAA 3660

Qy 3661 TGGCGCCCGGAGAAAGCATCTCTCTGCCAGCTGCAACGGCCTTGTGTACGGCAACAAGCT 3720

Db 3661 TGGCGCCCGGAGAAAGCATCTCTCTGCCAGCTGCAACGGCCTTGTGTACGGCAACAAGCT 3720

Qy 3721 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAA 3780

Db 3721 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAA 3780

Qy 3781 CTACATTAAGAGGATCTTCCCTCTTGGAAATGTCAACAAATCTCTAGAGCTGAGGAATAA 3840

Db 3781 CTACATTAAGAGGATCTTCCCTCTTGGAAATGTCAACAAATCTCTAGAGCTGAGGAATAA 3840

Qy 3841 AGATTTAGACATAGTCAACAGTCCAGCACACAAATCTACCTTGGCCACAGACCCCATGAG 3900

Db 3841 AGATTTAGACATAGTCAACAGTCCAGCACACAAATCTACCTTGGCCACAGACCCCATGAG 3900

Qy 3901 TGGGSCCGTCTTCCCTTCTGACAGCAACAGCCGCGGTCTTTAAATCAAGTCCACTGT 3960

Db 3901 TGGGSCCGTCTTCCCTTCTGACAGCAACAGCCGCGGTCTTTAAATCAAGTCCACTGT 3960

Qy 3961 GGTGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGGGGGACAGGTGACCAGTGCT 4020

Db 3961 GGTGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGGGGGACAGGTGACCAGTGCT 4020

Qy 4021 CCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAACCACTCACCAA 4080

Db 4021 CCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAACCACTCACCAA 4080

Qy 4081 TCCAGGGGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT 4140

Db 4081 TCCAGGGGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT 4140

Qy 4141 CAGACGCATCGATCAGAAATGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATC 4200

Db 4141 CAGACGCATCGATCAGAAATGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATC 4200

Qy 4201 AGCCCGGCCACTCAGCTGTGATTTCTCATGGATATTTCCAGGTAAGACTGGAGTGGCC 4260

Db 4201 AGCCCGGCCACTCAGCTGTGATTTCTGTATGGATATTTCCAGGTAAGACTGGAGTGGCC 4260

QY 4261 CACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCTCTCGACAACAATGTGGT 4320  
Db 4261 CACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCTCTCGACAACAATGTGGT 4320  
QY 4321 CCTGCAAAATCTGTGAAAAACCAACAGGTGGCATTTGTCCCGGGAGGCCCATGCACCTGCCA 4380  
Db 4321 CCTGCAAAATCTGTGAAAAACCAACAGGTGGCATTTGTCCCGGGAGGCCCATGCACCTGCCA 4380  
QY 4381 GGTCCCTGGCATTTGACCACTTCTCTGCTAAGCAAGGTGGCCATCCACGCAACCTTGGAGTC 4440  
Db 4381 GGTCCCTGGCATTTGACCACTTCTCTGCTAAGCAAGGTGGCCATCCACGCAACCTTGGAGTC 4440  
QY 4441 AGCCACCGCTTTGGCTGTTTTCACACAATGGGTCTCTGTATATTGCTGAGACTGATGAGAA 4500  
Db 4441 AGCCACCGCTTTGGCTGTTTTCACACAATGGGTCTCTGTATATTGCTGAGACTGATGAGAA 4500  
QY 4501 AAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4560  
Db 4501 AAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGC 4560  
QY 4561 CCCAGTGGCTGTCACTGTAAAAATGATGCCAACTGTGATGTTTCTGGAGACGATGG 4620  
Db 4561 CCCAGTGGCTGTCACTGTAAAAATGATGCCAACTGTGATGTTTCTGGAGACGATGG 4620  
QY 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCACTTCCCTGGCTGTGTGCTGATGGGA 4680  
Db 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCACTTCCCTGGCTGTGTGCTGATGGGA 4680  
QY 4681 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCTTT 4740  
Db 4681 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCTTT 4740  
QY 4741 CCTCAACACCCAGAACATGTATGAGCTGTCTTACCAATTTGACAGGAGCTCTATCTGTT 4800  
Db 4741 CCTCAACACCCAGAACATGTATGAGCTGTCTTACCAATTTGACAGGAGCTCTATCTGTT 4800  
QY 4801 TGATACCACCGGAAGCACCTGTATACACCAAGCCTGCCACAGGAGACTACCTGTACAA 4860  
Db 4801 TGATACCACCGGAAGCACCTGTATACACCAAGCCTGCCACAGGAGACTACCTGTACAA 4860  
QY 4861 CTTTACCTACACTGGGACGGGACATCACACTCATCACAGACAAACAATGGCAACATGGT 4920  
Db 4861 CTTTACCTACACTGGGACGGGACATCACACTCATCACAGACAAACAATGGCAACATGGT 4920  
QY 4921 AAATGTCGCGGACACTCTACTGGGATGCCCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT 4980  
Db 4921 AAATGTCGCGGACACTCTACTGGGATGCCCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT 4980  
QY 4981 GTACTGGGTGACATACCATACCAAGTGCATCAAGAGTGTGACCAACAAGGACACGA 5040  
Db 4981 GTACTGGGTGACATACCATACCAAGTGCATCAAGAGTGTGACCAACAAGGACACGA 5040  
QY 5041 GTTGGCCATGATGACATACCATACCAAGTGCATCAAGAGTGTGACCAACAAGGACAA 5100  
Db 5041 GTTGGCCATGATGACATACCATACCAAGTGCATCAAGAGTGTGACCAACAAGGACAA 5100  
QY 5101 CGGATGGACAAATTTATGATACGACAGCTTTGGCCGCTTACCAAAATGTGACCTTCCC 5160  
Db 5101 CGGATGGACAAATTTATGATACGACAGCTTTGGCCGCTTACCAAAATGTGACCTTCCC 5160  
QY 5161 TACTGGCCAGTGACAGTTCCTGGAAGTATACAGACAGTTTCTGCTGCTCAGGCTTCTA 5220  
Db 5161 TACTGGCCAGTGACAGTTCCTGGAAGTATACAGACAGTTTCTGCTGCTCAGGCTTCTA 5220  
QY 5221 GACCTCCAGCAAGGATGATGTACCAATAACCAACCACTGTCTGCTCAGGCGCTTCTA 5280  
Db 5221 GACCTCCAGCAAGGATGATGTACCAATAACCAACCACTGTCTGCTCAGGCGCTTCTA 5280  
QY 5281 CACACTGCTGCAAGCAAGTCCGGAACAGCTACTACATCGGGCCGATGGCTCCTTGG 5340  
Db 5281 CACACTGCTGCAAGCAAGTCCGGAACAGCTACTACATCGGGCCGATGGCTCCTTGG 5340  
QY 5341 GCTGCTGCTGCGCAACGGCATGGAGGTGGCTGCAGACTGAGCCCCACTTGTGCTGG 5400

Db 5341 GCTGCTGCTGCGCAACGGCATGGAGGTGGCTGCAGACTGAGCCCCACTTGTGCTGG 5400  
QY 5401 CACCGTCAACCCACCGTGGGCAAGAGGAATGTACGCTGCCATCGACAACGGCCTCAA 5460  
Db 5401 CACCGTCAACCCACCGTGGGCAAGAGGAATGTACGCTGCCATCGACAACGGCCTCAA 5460  
QY 5461 CCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCG 5520  
Db 5461 CCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCG 5520  
QY 5521 CCGGCTGGGCTGCACAACCCGAAATCTCCTATCTCTGGACTTTGATCGCTAACACGCAC 5580  
Db 5521 CCGGCTGGGCTGCACAACCCGAAATCTCCTATCTCTGGACTTTGATCGCTAACACGCAC 5580  
QY 5581 AGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTGTACGACGCGGG 5640  
Db 5581 AGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTGTACGACGCGGG 5640  
QY 5641 GCGGCCCGAGCCTCTGTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCC 5700  
Db 5641 GCGGCCCGAGCCTCTGTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCC 5700  
QY 5701 TGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAGAATGGAATACGACA 5760  
Db 5701 TGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAGAATGGAATACGACA 5760  
QY 5761 GCGGGCCCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATACT 5820  
Db 5761 GCGGGCCCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATACT 5820  
QY 5821 AGAGAAGTCCATGGTGTCTACTACACAGCCAGAGGCGAGTATATCTTTGAGTTGACAA 5880  
Db 5821 AGAGAAGTCCATGGTGTCTACTACACAGCCAGAGGCGAGTATATCTTTGAGTTGACAA 5880  
QY 5881 GAATGACCGCCTCTCTTCTGTGACGATGCCCAACGCTGGCGGCGAGACACTAGAGACCAT 5940  
Db 5881 GAATGACCGCCTCTCTTCTGTGACGATGCCCAACGCTGGCGGCGAGACACTAGAGACCAT 5940  
QY 5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAT 6000  
Db 5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAT 6000  
QY 6001 ACAGGACTTCACTGAGGATGGGCACTCTTACACCTTCTACCTGGGCACTGGCCGAG 6060  
Db 6001 ACAGGACTTCACTGAGGATGGGCACTCTTACACCTTCTACCTGGGCACTGGCCGAG 6060  
QY 6061 GGTGATATACAAGTATGGCAACTGTCAAAAGCTGGCAGAGACGCTCTATGACACCAACAA 6120  
Db 6061 GGTGATATACAAGTATGGCAACTGTCAAAAGCTGGCAGAGACGCTCTATGACACCAACAA 6120  
QY 6121 GGTGAGTTTCACTATGACGAGACGGCAGGCGATGCTGAAGACCATCAACCTACAGAATGA 6180  
Db 6121 GGTGAGTTTCACTATGACGAGACGGCAGGCGATGCTGAAGACCATCAACCTACAGAATGA 6180  
QY 6181 GGGCTTCACTGCACCATCCGCTACCGTACAGATTGGGCCCCCTGATTGACCGACAGATCTT 6240  
Db 6181 GGGCTTCACTGCACCATCCGCTACCGTACAGATTGGGCCCCCTGATTGACCGACAGATCTT 6240  
QY 6241 CCGCTTCACTGAGGAAGGCTGTTCAACGCCCTTTTGTACTACAATGATGATCTCTATCG 6300  
Db 6241 CCGCTTCACTGAGGAAGGCTGTTCAACGCCCTTTTGTACTACAATGATGATCTCTATCG 6300  
QY 6301 CCGGTTGACGAGCATGAGGCTGTGATCAACGAGACCCCACTGCCCCATTTGATCTCTATCG 6360  
Db 6301 CCGGTTGACGAGCATGAGGCTGTGATCAACGAGACCCCACTGCCCCATTTGATCTCTATCG 6360  
QY 6361 CTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTATTACTATGA 6420  
Db 6361 CTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTATTACTATGA 6420  
QY 6421 CATTAAACGATCATCACACAGCTGTCTATGACCCACCAAGCATTTTGTGATGCATATGG 6480

Db 6421 CATTAAACCAGATCATCACCACAGCTGTCTATGACCCACACCAAGCAATTTTGATGATATGG 6480  
QY 6481 CAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGATGACCGTCCA 6540  
Db 6481 CAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGATGACCGTCCA 6540  
QY 6541 GTATGATAACATGGGCGAGTAGTGAAGAAGAGCTGAAGGTAGGACCCCTACGCCAATAC 6600  
Db 6541 GTATGATAACATGGGCGAGTAGTGAAGAAGAGCTGAAGGTAGGACCCCTACGCCAATAC 6600  
QY 6601 CACTCGCTACTCCTATGAGTATGATGCTGACGGCGAGCTGCAGACAGTCTCCATCAATGA 6660  
Db 6601 CACTCGCTACTCCTATGAGTATGATGCTGACGGCGAGCTGCAGACAGTCTCCATCAATGA 6660  
QY 6661 CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCATTTACTGAGCCCTGG 6720  
Db 6661 CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCATTTACTGAGCCCTGG 6720  
QY 6721 GAACAGTGCACGGCTCACACCACACTACGGTATGACATCCGCGACCGCATCACTCGGCTGG 6780  
Db 6721 GAACAGTGCACGGCTCACACCACACTACGGTATGACATCCGCGACCGCATCACTCGGCTGG 6780  
QY 6781 TGACGTGCAATACAAGATGGATGAGGATGGCTTCTTGAGGCAGCGGGCGGTGATATCTT 6840  
Db 6781 TGACGTGCAATACAAGATGGATGAGGATGGCTTCTTGAGGCAGCGGGCGGTGATATCTT 6840  
QY 6841 TGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTTACAACCGGCTGGCAGCTGGAGTGT 6900  
Db 6841 TGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTTACAACCGGCTGGCAGCTGGAGTGT 6900  
QY 6901 CAGGTACCGCTACGATGGCTGGGGCGGCGTGTCCAGCAAGAGCAGCCACAGCCACCA 6960  
Db 6901 CAGGTACCGCTACGATGGCTGGGGCGGCGTGTCCAGCAAGAGCAGCCACAGCCACCA 6960  
QY 6961 CCTGCAGTTCTTCTATGACAGCCTGACCAACCCCAAGGTCAACCCACCTGTACAACCA 7020  
Db 6961 CCTGCAGTTCTTCTATGACAGCCTGACCAACCCCAAGGTCAACCCACCTGTACAACCA 7020  
QY 7021 CTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGCAAGGACACCTTTGGCCATGGA 7080  
Db 7021 CTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGCAAGGACACCTTTGGCCATGGA 7080  
QY 7081 GCTGAGCAGTGGTATGAGTTTACATAGCTTTGCAAAACATCGGACCCCTCTTGTCTGT 7140  
Db 7081 GCTGAGCAGTGGTATGAGTTTACATAGCTTTGCAAAACATCGGACCCCTCTTGTCTGT 7140  
QY 7141 CTTTAGTGGAACAGGTTTGATGATCAAGCAATTCCTGTACACAGCCTATGGGAGATCTA 7200  
Db 7141 CTTTAGTGGAACAGGTTTGATGATCAAGCAATTCCTGTACACAGCCTATGGGAGATCTA 7200  
QY 7201 CATGGATACCAACCCCAACTTTAGATCATATAGGCTACCATGGTGGCTCTATGATCC 7260  
Db 7201 CATGGATACCAACCCCAACTTTAGATCATATAGGCTACCATGGTGGCTCTATGATCC 7260  
QY 7261 ACTCACCAAGCTTGTCCACATGGGCGGCGAGATTATGATGTGCTGGCGGACGCTGGAC 7320  
Db 7261 ACTCACCAAGCTTGTCCACATGGGCGGCGAGATTATGATGTGCTGGCGGACGCTGGAC 7320  
QY 7321 TAGCCAGACCCAGCTGTGGAAGCACCTTTAGTAGCAGCAACGTCATGCTTTTAATCT 7380  
Db 7321 TAGCCAGACCCAGCTGTGGAAGCACCTTTAGTAGCAGCAACGTCATGCTTTTAATCT 7380  
QY 7381 CTATATGTTCAAAAACCAACCCCATCAGCAACTCCAGGACATCAAGTCTTCATGAC 7440  
Db 7381 CTATATGTTCAAAAACCAACCCCATCAGCAACTCCAGGACATCAAGTCTTCATGAC 7440  
QY 7441 AGATGTTAACAGCTGGCTGCTACCTTTGGATTCCAGCTACACAACGTCCTCTGTTA 7500  
Db 7441 AGATGTTAACAGCTGGCTGCTACCTTTGGATTCCAGCTACACAACGTCCTCTGTTA 7500  
QY 7501 TCCCAAACCCAGACATGGATGCCATGGAACCCCTCTACGAGCTCATCCACACAGATGAA 7560  
Db 7501 TCCCAAACCCAGACATGGATGCCATGGAACCCCTCTACGAGCTCATCCACACAGATGAA 7560

QY 7561 AACGCAGGAGTGGACAAACAGCAAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAGCA 7620  
Db 7561 AACGCAGGAGTGGACAAACAGCAAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAGCA 7620  
QY 7621 GGTCAAGGCCCTTTGTACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAG 7680  
Db 7621 GGTCAAGGCCCTTTGTACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAG 7680  
QY 7681 CTGSCAGCAGGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGG 7740  
Db 7681 CTGSCAGCAGGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGG 7740  
QY 7741 GGTCAAGTTTGCCTTGAAGSATGGCCGAGTGACCAAGACATCATCAGTGTGGCCAATGA 7800  
Db 7741 GGTCAAGTTTGCCTTGAAGSATGGCCGAGTGACCAAGACATCATCAGTGTGGCCAATGA 7800  
QY 7801 GGATGGCGGAAGGGTTGCTGCCATCTTGAAACCATGCCCCACTACCTAGAGAACCCTGCATT 7860  
Db 7801 GGATGGCGGAAGGGTTGCTGCCATCTTGAAACCATGCCCCACTACCTAGAGAACCCTGCATT 7860  
QY 7861 CACCATTGATGGGTGGATACCCATTACTTTGTGAAACCAAGGACCTTCAGAAAGGTGACCT 7920  
Db 7861 CACCATTGATGGGTGGATACCCATTACTTTGTGAAACCAAGGACCTTCAGAAAGGTGACCT 7920  
QY 7921 GGCCATCCTGGGCCTCAGTGGGGGGCGGCGAACCCCTGGAGAAATGGGTCAACGTCACTGT 7980  
Db 7921 GGCCATCCTGGGCCTCAGTGGGGGGCGGCGAACCCCTGGAGAAATGGGTCAACGTCACTGT 7980  
QY 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCA 8040  
Db 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCA 8040  
QY 8041 GTACGGGGCACTGTGCTTGAACACACAGCTACGGGCAACGTTGGATGAGGAGAAAGCACG 8100  
Db 8041 GTACGGGGCACTGTGCTTGAACACACAGCTACGGGCAACGTTGGATGAGGAGAAAGCACG 8100  
QY 8101 GGTCTGGAGCTGGCCCGCAGAGAGCCGTGCGCCAGCGTGGCCCGCAGCAGAG 8160  
Db 8101 GGTCTGGAGCTGGCCCGCAGAGAGCCGTGCGCCAGCGTGGCCCGCAGCAGAG 8160  
QY 8161 ACTCGGGAAGGGAGGAGGAGCCCTCGGGCCTGGACAGGGGGAGAGCAGAGTGCT 8220  
Db 8161 ACTCGGGAAGGGAGGAGGAGCCCTCGGGCCTGGACAGGGGGAGAGCAGAGTGCT 8220  
QY 8221 GAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8280  
Db 8221 GAGCACAGGGCGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8280  
QY 8281 AGAACTGTGACAGCAGCGCCCAACATCCACTTCATGACAGAGCGAGATGGGCCGAG 8340  
Db 8281 AGAACTGTGACAGCAGCGCCCAACATCCACTTCATGACAGAGCGAGATGGGCCGAG 8340  
QY 8341 GTGACAGAGAGGAC 8354  
Db 8341 GTGACAGAGAGGAC 8354

RESULT 2  
US-10-029-020-13  
; Sequence 13, Application US/10029020  
; Publication No. US2004003971A1  
; GENERAL INFORMATION:  
; APPLICANT: Gargolli et al.  
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-225  
; CURRENT APPLICATION NUMBER: US/10/029,020  
; CURRENT FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/256,704  
; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 60/311,590  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/257,314



Db 1801 ||||| CCGGACTGTGGCAGAGCTCCTGCCCGTCTGTAGCGGAAATGGCCAATACATGAA 1860

Qy 1861 AGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGTGAGTGCATGTGCCACCAACCA 1920

Db 1861 AGGCAGATGCTTGTGCCACAGTGGCTGGAAGGCGTGAGTGCATGTGCCACCAACCA 1920

Qy 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCACGGGCACCTGCAATGCAA 1980

Db 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCACGGGCACCTGCAATGCAA 1980

Qy 1981 CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCAACATGTTTCAGG 2040

Db 1981 CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCAACATGTTTCAGG 2040

Qy 2041 CCGGGGTGTCTGCGTGAAGCGCAATGCCATTGCTTTGTGGATGGGGAGGCACCAACTG 2100

Db 2041 CCGGGGTGTCTGCGTGAAGCGCAATGCCATTGCTTTGTGGATGGGGAGGCACCAACTG 2100

Qy 2101 CGAGACCCCCAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAAACCTTCTCTCCCGGA 2160

Db 2101 CGAGACCCCCAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAAACCTTCTCTCCCGGA 2160

Qy 2161 CACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG 2220

Db 2161 CACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG 2220

Qy 2221 TGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGG 2280

Db 2221 TGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGG 2280

Qy 2281 CTGGATGGGGGCAGCCTGCGACCAAGCGGGCCTGCCACCCGCGTGTGCCGAGCATGGGAC 2340

Db 2281 CTGGATGGGGGCAGCCTGCGACCAAGCGGGCCTGCCACCCGCGTGTGCCGAGCATGGGAC 2340

Qy 2341 CTGCCGCGACGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGC 2400

Db 2341 CTGCCGCGACGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGC 2400

Qy 2401 TCACTATCTGGATAGGTAAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAG 2460

Db 2401 TCACTATCTGGATAGGTAAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAG 2460

Qy 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGG 2520

Db 2461 ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGG 2520

Qy 2521 CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCT 2580

Db 2521 CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCT 2580

Qy 2581 GGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640

Db 2581 GGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640

Qy 2641 CCTTGGCTCCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGCA 2700

Db 2641 CCTTGGCTCCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGCA 2700

Qy 2701 GAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGACAGGTCCTGTGTACAGCA 2760

Db 2701 GAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGACAGGTCCTGTGTACAGCA 2760

Qy 2761 AATCCCCGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTTCGTGGCCAAAGTGAT 2820

Db 2761 AATCCCCGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTTCGTGGCCAAAGTGAT 2820

Qy 2821 GACATCAGATGGAACCCCTGGTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTT 2880

Db 2821 GACATCAGATGGAACCCCTGGTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTT 2880

Qy 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC 2940

Db 2881 TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC 2940

Qy 2941 CATCATCCTGCGGTTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCC 3000

Db 2941 CATCATCCTGCGGTTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCC 3000

Qy 3001 ATGGGATCGCTTCTTTTGTATGGAACCAATCATCATGAGACATGAGGAGAAATGAGATTCC 3060

Db 3001 ATGGGATCGCTTCTTTTGTATGGAACCAATCATCATGAGACATGAGGAGAAATGAGATTCC 3060

Qy 3061 CAGCTGTGACCTGAGCAATTTTGGCCGCCCCCAACCCAGTCGTCTCTCCATCCCCACTGAC 3120

Db 3061 CAGCTGTGACCTGAGCAATTTTGGCCGCCCCCAACCCAGTCGTCTCTCCATCCCCACTGAC 3120

Qy 3121 GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCCAATTTGTGCCGGAATTCAGGCTTTGCA 3180

Db 3121 GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCCAATTTGTGCCGGAATTCAGGCTTTGCA 3180

Qy 3181 GGAGGAAATCTCTATCTCTGGCTCAAGATGAGGTCAGCTACCTGAGCAGCCGACCCCC 3240

Db 3181 GGAGGAAATCTCTATCTCTGGCTCAAGATGAGGTCAGCTACCTGAGCAGCCGACCCCC 3240

Qy 3241 TGGCTACAAAATCTGTCTGAGGATCAGCCTCACCCACCGACCATCCCCCTTCAACCTCAT 3300

Db 3241 TGGCTACAAAATCTGTCTGAGGATCAGCCTCACCCACCGACCATCCCCCTTCAACCTCAT 3300

Qy 3301 GAAGTGCACCTCATGTTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGC 3360

Db 3301 GAAGTGCACCTCATGTTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTCGCTGCAGC 3360

Qy 3361 CCCAGACCTGTCTCTATTATTTCAATTTGGGACAAGACAGACGTCTACAACCAGAAAGTGTT 3420

Db 3361 CCCAGACCTGTCTCTATTATTTCAATTTGGGACAAGACAGACGTCTACAACCAGAAAGTGTT 3420

Qy 3421 TGGGCTTTCAGAAAGCCCTTTGTTCCGTGGTTATGAATATGAATCCTGCCCAGATCTAAT 3480

Db 3421 TGGGCTTTCAGAAAGCCCTTTGTTCCGTGGTTATGAATATGAATCCTGCCCAGATCTAAT 3480

Qy 3481 CCTGTGGGAAAAAGAACACAGTCTGCAGGGCTATGAAATTGACGCGTCCAAGCTTGG 3540

Db 3481 CCTGTGGGAAAAAGAACACAGTCTGCAGGGCTATGAAATTGACGCGTCCAAGCTTGG 3540

Qy 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTGCACAAAGG 3600

Db 3541 AGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTGCACAAAGG 3600

Qy 3601 GAATGGGAGAACCAAGTTTGTCTTCAGCAGCCTCCTGTCAATTTGGGAGCATCATGGGCAA 3660

Db 3601 GAATGGGAGAACCAAGTTTGTCTTCAGCAGCCTCCTGTCAATTTGGGAGCATCATGGGCAA 3660

Qy 3661 TGGGCGCGGAGAAAGCATCTCCTGCCCCAGCTGCAACGGCTTGTGACGGCAACAAGCT 3720

Db 3661 TGGGCGCGGAGAAAGCATCTCCTGCCCCAGCTGCAACGGCTTGTGACGGCAACAAGCT 3720

Qy 3721 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAA 3780

Db 3721 CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAA 3780

Qy 3781 CTACATTAGAAGGATCTTCCCCTCTGGAAATGTCAACCAATCCTTAGAGCTGAGGAATAA 3840

Db 3781 CTACATTAGAAGGATCTTCCCCTCTGGAAATGTCAACCAATCCTTAGAGCTGAGGAATAA 3840

Qy 3841 AGATTTTCAGACATAGTCAAGTCCAGCACACAAATACTACTCTGGCCACAGACCCCATGAG 3900

Db 3841 AGATTTTCAGACATAGTCAAGTCCAGCACACAAATACTACTCTGGCCACAGACCCCATGAG 3900

Qy 3901 TGGGGCCGTCTTCTTTCTGACAGCAACAGCCCGGGCTTTTAAATCAAGTCCACTGT 3960

Db 3901 TGGGGCCGTCTTCTTTCTGACAGCAACAGCCCGGGCTTTTAAATCAAGTCCACTGT 3960

Qy 3961 GGTGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAGTGCCT 4020

Db 3961 GGTGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAGTGCCT 4020

QY	4021	CCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACCAA	4081
Db	4021	CCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACCAA	4080
QY	4081	TCCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT	4140
Db	4081	TCCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT	4140
QY	4141	CAGACGCATCGATCAGAAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC	4200
Db	4141	CAGACGCATCGATCAGAAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC	4200
QY	4201	AGCCCGGCCACTCAGCTGTGATTCTGTCAATGGATATTCCCAAGGTAAAGACTGGAGTGGCC	4260
Db	4201	AGCCCGGCCACTCAGCTGTGATTCTGTCAATGGATATTCCCAAGGTAAAGACTGGAGTGGCC	4260
QY	4261	CACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCTCGACAACAATCTGGT	4320
Db	4261	CACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCTCGACAACAATCTGGT	4320
QY	4321	CCTGCAAAATCTCTGAAAACCAACAGGTGCGCATTTGCGCCGGAGGCCCATGCACTGCCA	4380
Db	4321	CCTGCAAAATCTCTGAAAACCAACAGGTGCGCATTTGCGCCGGAGGCCCATGCACTGCCA	4380
QY	4381	GGTCCCTGGCATTGACCACTTCTGTAAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC	4440
Db	4381	GGTCCCTGGCATTGACCACTTCTGTAAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC	4440
QY	4441	AGCCACCCTTTGGCTGTTTACACAATGGGTCCTGTATATTGCTGAGACTGATGAGAA	4500
Db	4441	AGCCACCCTTTGGCTGTTTACACAATGGGTCCTGTATATTGCTGAGACTGATGAGAA	4500
QY	4501	AAAGATCAACCCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGC	4560
Db	4501	AAAGATCAACCCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGC	4560
QY	4561	CCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTTCTGGAGACGATGG	4620
Db	4561	CCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTTCTGGAGACGATGG	4620
QY	4621	TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTTGGCTGTGTGCTGATGGGA	4680
Db	4621	TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTTGGCTGTGTGCTGATGGGA	4680
QY	4681	GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGTTTTATCCGGAAGAACBAGCCTTT	4740
Db	4681	GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGTTTTATCCGGAAGAACBAGCCTTT	4740
QY	4741	CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAGGAGCTCTATCTGTT	4800
Db	4741	CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAGGAGCTCTATCTGTT	4800
QY	4801	TGATACCACCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAA	4860
Db	4801	TGATACCACCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAA	4860
QY	4861	CTTCACTACACTGGGGACGGCGACATCACACTCATCAGACAACAATGGCAACATGGT	4920
Db	4861	CTTCACTACACTGGGGACGGCGACATCACACTCATCAGACAACAATGGCAACATGGT	4920
QY	4921	AAATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT	4980
Db	4921	AAATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT	4980
QY	4981	GTA CTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGACCAACAAAGCACCGA	5040
Db	4981	GTA CTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGACCAACAAAGCACCGA	5040
QY	5041	GTTGGCCATGATGACATACCATGGCAATTCCGGCTTCTGGCAACCAAAAGCAATGAAAA	5100
Db	5041	GTTGGCCATGATGACATACCATGGCAATTCCGGCTTCTGGCAACCAAAAGCAATGAAAA	5100

QY	5101	CGGATGGACAAACATTTTATGAGTACGACAGCTTTGGCCGCCTGACAAATGTGACCTTCCC	5160
Db	5101		5160
QY	5161	TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCCAGTGCATGTCCAGGTAGA	5220
Db	5161		5220
QY	5221	GACCTCCAGCAAGGATGTACCCATAACACACCAACCTGTCTGCCTCAGGCGCCTTCTA	5280
Db	5221		5280
QY	5281	CACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCCTTGCG	5340
Db	5281		5340
QY	5341	GCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTGTCTGGCTGG	5400
Db	5341		5400
QY	5401	CACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACAGCTGCCCATCGACAACGGCCTCAA	5460
Db	5401		5460
QY	5461	CCTGGTGGAGTGGCGCCAGCGCAAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGGCG	5520
Db	5461		5520
QY	5521	CCGGCTGCGGTGCACAAACCGAAATCTCTATCTCTGGACTTTTGATCGCGTAACACGCAC	5580
Db	5521		5580
QY	5581	AGAGAAGATCTATGATGACCACCGCAAGTTCACTTCCGATTTCTGTACGACCAGGCGGG	5640
Db	5581		5640
QY	5641	GCGGCCAGCCTCTGGTCAACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC	5700
Db	5641		5700
QY	5701	TGGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCA	5760
Db	5701		5760
QY	5761	GGCGGGCCGCATCACATCCAGGATCTTCGCTCATGGGAAGACATGGAGCTACACATACTT	5820
Db	5761		5820
QY	5821	AGAGAAGTCCATGGTGCTGCTACTACAGACCCAGAGGCAGTATATCTTTGAGTTCGACAA	5880
Db	5821		5880
QY	5881	GAATGACCCCTCTCTTCTGTGACGATGCCCAACGTGGCGCGGCAGACACTAGAGACCAT	5940
Db	5881		5940
QY	5941	CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAT	6000
Db	5941		6000
QY	6001	ACAGGACTTCACTGAGGATGGGCACTCTCTTACACACCTTCTACCTGGGCACTGGCCCGCAG	6060
Db	6001		6060
QY	6061	GGTGATATACAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAA	6120
Db	6061		6120
QY	6121	GGTCAGTTTACCTTATGACGAGACGGCAGGCATGCTGAAGAACCTCAACCTACAGAATGA	6180
Db	6121		6180
QY	6181	GGGCTTACCTGCACCATCCGCTACCGTACAGATTGGGCCCTTGATTGACCGACAGATCTT	6240

Db 6181 GGGCTTACCTGCACCATCCGTTACCGTACAGATTGGGCCCTGATTGACCGACAGATCTT 6240  
QY 6241 CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTGTACTAACTATGACAAACAGCTT 6300  
Db 6241 CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTGTACTAACTATGACAAACAGCTT 6300  
QY 6301 CCGGGTGACCAAGCATGAGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCG 6360  
Db 6301 CCGGGTGACCAAGCATGAGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTCTATCG 6360  
QY 6361 CTATGATGATGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGA 6420  
Db 6361 CTATGATGATGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGA 6420  
QY 6421 CATTAACAGATCATCAACACAGCTGTCAATGACCCCAAGCATTTTGTATGTCATATGG 6480  
Db 6421 CATTAACAGATCATCAACACAGCTGTCAATGACCCCAAGCATTTTGTATGTCATATGG 6480  
QY 6481 CAGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGTCACTGATGATGACCGTCCA 6540  
Db 6481 CAGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGTCACTGATGATGACCGTCCA 6540  
QY 6541 GTATGATAACATGGGGGAGTAGTGAAGAAGGAGTGAAGGCTAGGACCCCTACGCCAATAC 6600  
Db 6541 GTATGATAACATGGGGGAGTAGTGAAGAAGGAGTGAAGGCTAGGACCCCTACGCCAATAC 6600  
QY 6601 CACTCGCTACTCCTATGAGTATGATGCTGACGCGCCAGCTGCAGACAGTCTCCATCAATGA 6660  
Db 6601 CACTCGCTACTCCTATGAGTATGATGCTGACGCGCCAGCTGCAGACAGTCTCCATCAATGA 6660  
QY 6661 CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG 6720  
Db 6661 CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG 6720  
QY 6721 GAACAGTGCACGGCTCACACCACCTACGGTATGACATCCGCGACCGCATCACTCGGCTGG 6780  
Db 6721 GAACAGTGCACGGCTCACACCACCTACGGTATGACATCCGCGACCGCATCACTCGGCTGG 6780  
QY 6781 TGACGTGCAATACAAGATGGATGAGGATGGCTTCCCTGAGGCAGCGGGCGGTGATATCTT 6840  
Db 6781 TGACGTGCAATACAAGATGGATGAGGATGGCTTCCCTGAGGCAGCGGGCGGTGATATCTT 6840  
QY 6841 TGAGTACAACCTCAGCTGGCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGT 6900  
Db 6841 TGAGTACAACCTCAGCTGGCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGT 6900  
QY 6901 CAGGTACCGCTACGATGGCTGGGGCGGCGCTGTCCAGCAAGAGCAGCCACCA 6960  
Db 6901 CAGGTACCGCTACGATGGCTGGGGCGGCGCTGTCCAGCAAGAGCAGCCACCA 6960  
QY 6961 CCTGCAGTTCTTCTATGCAGACCTGACCAACCCCAACCAAGGTCAACCCACCTGTACAACCA 7020  
Db 6961 CCTGCAGTTCTTCTATGCAGACCTGACCAACCCCAACCAAGGTCAACCCACCTGTACAACCA 7020  
QY 7021 CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGA 7080  
Db 7021 CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGA 7080  
QY 7081 GCTGAGCAGTGGTATGAGTTTACATAGCTTGTGACAAACATCGGGACCCCTCTTGTCTGT 7140  
Db 7081 GCTGAGCAGTGGTATGAGTTTACATAGCTTGTGACAAACATCGGGACCCCTCTTGTCTGT 7140  
QY 7141 CTTTAGTGAACAGGTTTGTATGATCAAGCAAATCCTGTACACAGCCTATGGGAGATCTA 7200  
Db 7141 CTTTAGTGAACAGGTTTGTATGATCAAGCAAATCCTGTACACAGCCTATGGGAGATCTA 7200  
QY 7201 CATGGATACCAACCCCACTTTCAGATCATATAGCTACCATGGTGGCCTCTATGATCC 7260  
Db 7201 CATGGATACCAACCCCACTTTCAGATCATATAGCTACCATGGTGGCCTCTATGATCC 7260  
QY 7261 ACTCACCAAGCTTGTCCATGCGGGCGGAGATTATGATGTGCTGGCCGACCGCTGGAC 7320

Db 7261 ACTCACCAAGCTTGTCCACATGGGGCGGAGATTATGATGTGTGGCCGGACGCTGGAC 7320  
QY 7321 TAGCCCAAGACCAAGAGCTGTGGAAGACACCTTAGTAGCAGCAACGTCATGCCTTTTAATCT 7380  
Db 7321 TAGCCCAAGACCAAGAGCTGTGGAAGACACCTTAGTAGCAGCAACGTCATGCCTTTTAATCT 7380  
QY 7381 CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCCAGGACATCAAGTGTTCATGAC 7440  
Db 7381 CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCCAGGACATCAAGTGTTCATGAC 7440  
QY 7441 AGATGTTAAACAGCTGGCTGCTCCTCACCTTTGGATTCCAGCTACACAACGTCATCCCTGGTTA 7500  
Db 7441 AGATGTTAAACAGCTGGCTGCTCCTCACCTTTGGATTCCAGCTACACAACGTCATCCCTGGTTA 7500  
QY 7501 TCCCAAAACCAAGACATGGATGCCATGGAACCCCTCCTACGAGCTCATCCACACAGATGAA 7560  
Db 7501 TCCCAAAACCAAGACATGGATGCCATGGAACCCCTCCTACGAGCTCATCCACACAGATGAA 7560  
QY 7561 AACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA 7620  
Db 7561 AACGCAGGAGTGGGACAAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA 7620  
QY 7621 GCTCAAGGCCCTTTGTACCTTGAACGGTTTGAACAGCTCTATGGCTCCACAATCACAG 7680  
Db 7621 GCTCAAGGCCCTTTGTACCTTGAACGGTTTGAACAGCTCTATGGCTCCACAATCACAG 7680  
QY 7681 CTGCCAGAGGGCTCCAAAAGACCAGAAAGTTTGCATCCAGGGCTCAGTCTTTGGCAAGGG 7740  
Db 7681 CTGCCAGAGGGCTCCAAAAGACCAGAAAGTTTGCATCCAGGGCTCAGTCTTTGGCAAGGG 7740  
QY 7741 GGTCAGTTTGCCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTGGCCAATGA 7800  
Db 7741 GGTCAGTTTGCCTTGAAGGATGGCCGAGTGACCAACAGACATCATCAGTGTGGCCAATGA 7800  
QY 7801 GGATGGCGAAGGGTTGCTGCCATCTTGAACCATGCCCACTACTAGAGAACCTGCACCTT 7860  
Db 7801 GGATGGCGAAGGGTTGCTGCCATCTTGAACCATGCCCACTACTAGAGAACCTGCACCTT 7860  
QY 7861 CACCAATTGATGGGTGGATACCCATTACTTTGTGAAACCAAGACCTTCAAGAGGTGACCT 7920  
Db 7861 CACCAATTGATGGGTGGATACCCATTACTTTGTGAAACCAAGACCTTCAAGAGGTGACCT 7920  
QY 7921 GGCCATCCTGGCCCTCAGTGGGGCGGCGAACCCTGGAGAATGGGTCAACGTCACCTGT 7980  
Db 7921 GGCCATCCTGGCCCTCAGTGGGGCGGCGAACCCTGGAGAATGGGTCAACGTCACCTGT 7980  
QY 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8040  
Db 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8040  
QY 8041 GTACGGGCACTGTGCTTTGAACACACGCTACGGGACAAAGCTTGGATGAGGAGAAAGGACG 8100  
Db 8041 GTACGGGCACTGTGCTTTGAACACACGCTACGGGACAAAGCTTGGATGAGGAGAAAGGACG 8100  
QY 8101 GGTCTGGAGCTGGCCCGGACAGAGCCGTGCGCCAAAGCTGGGCCCGGAGCAGAGAG 8160  
Db 8101 GGTCTGGAGCTGGCCCGGACAGAGCCGTGCGCCAAAGCTGGGCCCGGAGCAGAGAG 8160  
QY 8161 ACTGCGGGAAGGGAGGAAGGCTGCGGGCCTGGACAGAGGGGGAGAGCAGAGTGCT 8220  
Db 8161 ACTGCGGGAAGGGAGGAAGGCTGCGGGCCTGGACAGAGGGGGAGAGCAGAGTGCT 8220  
QY 8221 GAGCACAGGGCGGGTGCAAGGTCAGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8280  
Db 8221 GAGCACAGGGCGGGTGCAAGGTCAGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8280  
QY 8281 AGAATGTTCAGACAGCGCCCAACAACATCCACTTTCATGACACAGCGAGATGGGCCGAG 8340  
Db 8281 AGAATGTTCAGACAGCGCCCAACAACATCCACTTTCATGACACAGCGAGATGGGCCGAG 8340  
QY 8341 GTGACAGAGAGGAC 8354  
Db 8341 GTGACAGAGAGGAC 8354



QY 1535 AGGCTCCTAAACCCAGGAGCGCGAGCCCTAGAGGGACCCCGCGCCAGTCTCGGGAACT 1594  
Db |||||  
1501 AGGCTCCTAAACCCAGGAGCGCGAGCCCTAGAGGGACCCCGCGCCAGTCTCGGGAACT 1560  
QY 1595 GTGCCGCCCTCCAGCCATGAGACAGGCTTCCATCCAGTATTGGATTTCAGGAATCTGGCAC 1654  
Db |||||  
1561 GTGCCGCCCTCCAGCCATGAGACAGGCTTCCATCCAGTATTGGATTTCAGGAATCTGGCAC 1620  
QY 1655 TTGGCTTTTACAATGACGGAAAGGAGTCAGAAAGTGGTTTCCTTTCTCACCACCTGCCATT 1714  
Db |||||  
1621 TTGGCTTTTACAATGACGGAAAGGAGTCAGAAAGTGGTTTCCTTTCTCTCCCC-----A 1674  
QY 1715 GAGTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGACTGCACTCTCTGGGACC 1774  
Db |||||  
1675 GAGTCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGACTGCACTCTCTGGGACC 1734  
QY 1775 TGCACCTGCTTCTCTGGGTTTCTCTGGGCCCCGACTGTGGCAGAGCCTCCTGCCCGGTGCTC 1834  
Db |||||  
1735 TGCACCTGCTTCTCTGGGTTTCTCTGGGCCCCGACTGTGGCAGAGCCTCCTGCCCGGTGCTC 1794  
QY 1835 TGTAGCGGAAATGGCCAAATACATGAAGGCGAGATGCTGTGCCACAGTGGCTGCGAAAGGC 1894  
Db |||||  
1795 TGTAGCGGAAATGGCCAAATACATGAAGGCGAGATGCTGTGCCACAGTGGCTGCGAAAGGC 1854  
QY 1895 GCTGAGTCGGATGTGCCCAACCAACCAAGTGATCGATGTGGCCCTGCAGCAACCATGGCAC 1954  
Db |||||  
1855 GCTGAGTCGGATGTGCCCAACCAACCAAGTGATCGATGTGGCCCTGCAGCAACCATGGCAC 1914  
QY 1955 TGCATACGGGCACCTGCACTCTGCAACCCCTGGCTACAAAGGGCGAGAGCTGTGAGGAA 2014  
Db |||||  
1915 TGCATACGGGCACCTGCACTCTGCAACCCCTGGCTACAAAGGGCGAGAGCTGTGAGGAA--- 1971  
QY 2015 GACTGCATGGACCCCACTGTTTCAGGCCGGGTGCTGCTGAGAGCGGAATGCCATTGC 2074  
Db |||||  
1972 GACTGCATGGACCCCACTGTTTCAGGCCGGGTGCTGCTGAGAGCGGAATGCCACTGC 2031  
QY 2075 TTTGTGGGATGGGAGGACCAACTGCGAGACCCCGGGCTTGCAGCTGTAGACCAAGTGT 2134  
Db |||||  
2032 TCTGTGGGATGGGAGGACCAACTGCGAGACCCCGGGCTTGCAGCTGTAGACCAAGTGT 2091  
QY 2135 TCAGGCCACGGAACTTCTCTCCCGGACACCGGGCTTTCAGCTGTGACCCCAAGCTGGACT 2194  
Db |||||  
2092 TCAGGCCACGGAACTTCTCTCCCGGACACCGGGCTTTCAGCTGTGACCCCAAGCTGGACT 2151  
QY 2195 GGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTA 2254  
Db |||||  
2152 GGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTA 2211  
QY 2255 GGGGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGGACGCTGCGACAGCGGGCCTGC 2314  
Db |||||  
2212 GGGGGCACCTGCCGCTGCGAGGATGGCTGGATGGGGGACGCTGCGACAGCGGGCCTGC 2271  
QY 2315 CACCCGCGCTGTGCCGAGCATGGGACCTGCCGCGACGGCAAGTGGAGTGCAGCCCTGGC 2374  
Db |||||  
2272 CACCCGCGCTGTGCCGAGCATGGGACCTGCCGCGACGGCAAGTGGAGTGCAGCCCTGGC 2331  
QY 2375 TGGAAATGGCGAACACTGCACCAATC-----GCTCACTATCTGGATAGGTAGTTAA---- 2424  
Db |||||  
2332 TGGAAATGGCGAACACTGCACCAATCTCCCTAGCTCACTATCTGGATAGGTAGTTAACTT 2391  
QY 2425 --AGAGGGTTGCCCTGGGTTGTGCAATGGCAAGCGCAGATGTACCTTAGACCTTGAATGGT 2482  
Db |||||  
2392 TCAGAGGGTTGCCCTGGGTTGTCAATGGCAAGCGCAGATGTACCTTAGACCTTGAATGGT 2451  
QY 2483 TGGCACTGCGTCTGCCAGCTGGGCTGGAGAGAGCTGGCTGTGACACTTCCATGGAGACT 2542  
Db |||||  
2452 TGGCACTGCGTCTGCCAGCTGGGCTGGAGAGAGCTGGCTGTGACACTTCCATGGAGACT 2511  
QY 2543 GCCTGCGGTGACAGCAAGACAAATGATGGAGATGGCCTGGTGGACTGCACTGGACCCCTGAC 2602  
Db |||||  
2512 GCCTGCGGTGACAGCAAGACAAATGATGGAGATGGCCTGGTGGACTGCACTGGACCCCTGAC 2571

QY 2603 TGCTGCCTCCAGCCCTGTGTGCCATATCAAAACCCGCTGTGCCTTGGCTCCCTAACCCCTCTG 2662  
Db |||||  
2572 TGCTGCCTCCAGCCCTGTGTGCCATATCAAAACCCGCTGTGCCTTGGCTCCCTAACCCCTCTG 2631  
QY 2663 GACATCATCCAGGAGACACAGGTCCCTGTGTACAGCAGAACCTACACTCCTTCTATGAC 2722  
Db |||||  
2632 GACATCATCCAGGAGACACAGGTCCCTGTGTACAGCAGAACCTACACTCCTTCTATGAC 2691  
QY 2723 CGCATCAAGTTCTCGTGGGACGAGCAGCAGCACATATCCCGGGGAGAACCCCTTT 2782  
Db |||||  
2692 CGCATCAAGTTCTCGTGGGACGAGCAGCAGCACATATCCCGGGGAGAACCCCTTT 2751  
QY 2783 GATGGAGGCGATGCTTGTGTATTTCGTGGCCAAAGTGATGACATCAGATGGAACCCCTCTG 2842  
Db |||||  
2752 GATGGAGGCGATGCTTGTGTATTTCGTGGCCAAAGTGATGACATCAGATGGAACCCCTCTG 2811  
QY 2843 GTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATATACAATCAGCAGGCAA 2902  
Db |||||  
2812 GTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATATACAATCAGCAGGCAA 2871  
QY 2903 GATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCTCGGGTTCGAGCGG 2962  
Db |||||  
2872 GATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCTCGGGTTCGAGCGG 2931  
QY 2963 GCACCTTTCATCACACAGGAGCACACCTGTGGCTGCCATGGGATCGCTTCTTTGTGATG 3022  
Db |||||  
2932 GCACCTTTCATCACACAGGAGCACACCTGTGGTGCCATGGGATCGCTTCTTTGTGATG 2991  
QY 3023 GAAACCATCATCATGAGACATGAGGAGAAATGAGATTCCAGCTGTGACCTGAGCAATTT 3082  
Db |||||  
2992 GAAACCATCATCATGAGACATGAGGAGAAATGAGATTCCAGCTGTGACCTGAGCAATTT 3051  
QY 3083 GCCCGCCCCAACCCAGTCTCTCCATCCCCACTGACGTCTTGGCCAGCTCCTGTGCA 3142  
Db |||||  
3052 GCCCGCCCCAACCCAGTCTCTCCATCCCCACTGACGTCTTGGCCAGCTCCTGTGCA 3111  
QY 3143 GAGAAAGGCCCATTTGTGCCGAAATTCAGGCTTTTGAGGAGGAAATCTCTATCTCTGGC 3202  
Db |||||  
3112 GAGAAAGGCCCATTTGTGCCGAAATTCAGGCTTTTGAGGAGGAAATCTCTATCTCTGGC 3171  
QY 3203 TGCAAGATGAGGCTGAGCTACTGAGCAGCGCGACCCCTGGCTACAAATCTGTCTCTGAGG 3262  
Db |||||  
3172 TGCAAGATGAGGCTGAGCTACTGAGCAGCGCGACCCCTGGCTACAAATCTGTCTCTGAGG 3231  
QY 3263 ATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGCTGAGCG 3322  
Db |||||  
3232 ATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGCTGAGCG 3291  
QY 3323 GTGAGGGCCCGCTCTTTCAGGAAGTGGTTGCTGTCAGCCCCCAGACCTGTCTCTATTATTC 3382  
Db |||||  
3292 GTGAGGGCCCGCTCTTTCAGGAAGTGGTTGCTGTCAGCCCCCAGACCTGTCTCTATTATTC 3351  
QY 3383 ATTTGGGACAAGACAGACAGCTCTACAACCAAGAGGTGTTTGGGCTTTTCAGAAAGCCTTTGTT 3442  
Db |||||  
3352 ATTTGGGACAAGACAGACAGCTCTACAACCAAGAGGTGTTTGGGCTTTTCAGAAAGCCTTTGTT 3411  
QY 3443 TCCGTGGGTTATGAATATGAATCCTGCCCCAGATCTAATCTCTGTGGGAAAAAGAACACA 3502  
Db |||||  
3412 TCCGTGGGTTATGAATATGAATCCTGCCCCAGATCTAATCTCTGTGGGAAAAAGAACACA 3471  
QY 3503 GTGCTGAGGGCTATGAATATGAACCGCTCCAAGCTTGGAGGATGGAGCCTAGACAAACAT 3562  
Db |||||  
3472 GTGCTGAGGGCTATGAATATGAATGATGCTCCAAGCTTGGAGGATGGAGCCTAGACAAACAT 3531  
QY 3563 CATGCCCTCAACATTCAAAAGTGGT---ATCCTGCACAAAGGGAATGGGAGAACCAAGTTT 3619  
Db |||||  
3532 CATGCCCTCAACATTCAAAAGTGGTGGCATCTCTGCACAAAGGGAATGGGAGAACCAAGTTT 3591  
QY 3620 GTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATCATGGGCAATGGGCGCGGAGAGCATC 3679  
Db |||||  
3592 GTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATCATGGGCAATGGGCGCGGAGAGCATC 3651  
QY 3680 TCCTGCCCCAGCTGCAACGGCCTTGCTGACGGGCAACAAGCTCCTTGCCCCCAGTGGCCCTC 3739

Db 3652 TCCTGCCCCAGCTGCAACGGCCCTTGTGACGGCAACAAGCTCCTGGCCCCAGTGGCCCTC 3711  
QY 3740 ACCTGTGGTCTCTGACGGGAGCCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTC 3799  
Db 3712 ACCTGTGGTCTCTGACGGGAGCCCTCTATGTGGGTGATTTCAACTACATTAGAAGGATCTTC 3771  
QY 3800 CCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGG-----AATAAAGATTTCAACAT 3853  
Db 3772 CCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGTCAGAAATAAAGATTTCAACAT 3831  
QY 3854 AGTCACAGTCACAGCACACAAATACCTTGCCACAGACCCCATGAGTGGGCCCTCTTC 3913  
Db 3832 AGTCACAGTCACAGCACACAAATACCTTGCCACAGACCCCATGAGTGGGCCCTCTTC 3891  
QY 3914 CTTTCTGACAGCAACAGCCGGGCTCTTAAATCAAGTCACACTGTGGTGGTGAAGGAC 3973  
Db 3892 CTTTCTGACAGCAACAGCCGGGCTCTTAAATCAAGTCACACTGTGGTGGTGAAGGAC 3951  
QY 3974 CTTTGTCAAGAACTCTGAGGTGTTGCGGGACAGGTGACCACTGCCCTCCCTTTGATGAC 4033  
Db 3952 CTTTGTCAAGAACTCTGAGGTGTTGCGGGACAGGTGACCACTGCCCTCCCTTTGATGAC 4011  
QY 4034 ACTCGCTGCGGGATGTTGGGAAGGCCACAGAAAGCCACACTCACCAATCCCAGGGT--- 4090  
Db 4012 ACTCGCTGCGGGATGTTGGGAAGGCCACAGAAAGCCACACTCACCAATCCCAGGGTCCC 4071  
QY 4091 -----ATTACAGTGGACAAAGTTTGGGTGATCTACTTCGTGGATGGCACCATGATCAGA 4144  
Db 4072 CCAGGCATTTACAGTGGACAAAGTTTGGGTGATCTACTTCGTGGATGGCACCATGATCAGA 4131  
QY 4145 CGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCACATCAGCC 4204  
Db 4132 CGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCACATCAGCC 4191  
QY 4205 CGGCCACTCAGCTGTGATTTCTGTATGGATATTTCCAGGTAAGA-----CTGGAG 4255  
Db 4192 CGGCCACTCAGCTGTGATTTCTGTATGGATATTTCCAGGTAAGACAGGTTCACTTGGAG 4251  
QY 4256 TGGCCACAGACTTAGCCATCAACCCCAATGGACAACTCACTTTATGTCCTCGACAAAT 4315  
Db 4252 TGGCCACAGACTTAGCCATCAACCCCAATGGACAACTCACTTTATGTCCTCGACAAAT 4311  
QY 4316 GTGGTCTCTGAAATCTCTGAAACCCACAGGTGCGCATTTGTGCGGGAGGCCCATGAC 4375  
Db 4312 GTGGTCTCTGAAATCTCTGAAACCCACAGGTGCGCATTTGTGCGGGAGGCCCATGAC 4371  
QY 4376 TGGCAGGTCCCTGGCATTGACCACTTCTGTAAGCAAGTGGCCATCCACGCAACCTG 4435  
Db 4372 TGGCAGGTCCCTGGCATTGACCACTTCTGTAAGCAAGTGGCCATCCACGCAACCTG 4431  
QY 4436 GAGTCAGCCACCGCTTTGGCTGTTTACACAATGGGGTCTGTATATTGCTGAGACTGAT 4495  
Db 4432 GAGTCAGCCACCGCTTTGGCTGTTTACACAATGGGGTCTGTATATTGCTGAGACTGAT 4491  
QY 4496 GAGAAAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCT 4555  
Db 4492 GAGAAAAGATCAACCGCATCAGGAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCT 4551  
QY 4556 GGGGCCCCCAGTGGTGTGACTGTAAATAATGATGCCAAGTGTGATTTTCTGGAGAC 4615  
Db 4552 GGGGCCCCCAGTGGTGTGACTGTAAATAATGATGCCAAGTGTGATTTTCTGGAGAC 4611  
QY 4616 GATGTTATGCCAAGATGCAAGTTAAATAATGATGCCAAGTGTGATTTTCTGGAGAC 4675  
Db 4612 GATGTTATGCCAAGATGCAAGTTAAATAATGATGCCAAGTGTGATTTTCTGGAGAC 4671  
QY 4676 GGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAG 4735  
Db 4672 GGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAG 4731  
QY 4736 CCTTCTCAACACCCAGAACATGTATGAGCTGCTTTCACCAATTTGACGAGGAGCTCTAT 4795

Db 4732 CCTTCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAGGAGCTCTAT 4791  
QY 4796 CTGTTTGTATACACCGGCAAGACCTGTGTACACCCAAAGCCTGCCACAGGAGACTACCTG 4855  
Db 4792 CTGTTTGTATACACCGGCAAGACCTGTGTACACCCAAAGCCTGCCACAGGAGACTACCTG 4851  
QY 4856 TACAACTTCACTACACTGCGGACCGCGACATCACTCATCAAGACAAATGGCAAC 4915  
Db 4852 TACAACTTCACTACACTGCGGACCGCGACATCACTCATCAAGACAAATGGCAAC 4911  
QY 4916 ATGGTAAATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTCCAGATGGC 4975  
Db 4912 ATGGTAAATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTCCAGATGGC 4971  
QY 4976 CAGGTGTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACCAACAGGA 5035  
Db 4972 CAGGTGTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACCAACAGGA 5031  
QY 5036 CACGAGTTGCCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAAT 5095  
Db 5032 CACGAGTTGCCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAAT 5091  
QY 5096 GAAAAACGGATGGACAAATTTTATGAGTACGACAGCTTTGGCCGCTTGCACAAATGTGACC 5155  
Db 5092 GAAAAACGGATGGACAAATTTTATGAGTACGACAGCTTTGGCCGCTTGCACAAATGTGACC 5151  
QY 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAAGTGTATACAGACAGTTTCAAGTGCATGTCAG 5215  
Db 5152 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAAGTGTATACAGACAGTTTCAAGTGCATGTCAG 5211  
QY 5216 GTAGAGACCTCCAGCAAGGATGATGTACACCAATTAACCAACCACTGTCTGCCCTCAGGCGCC 5275  
Db 5212 GTAGAGACCTCCAGCAAGGATGATGTACCAATTAACCAACCACTGTCTGCCCTCAGGCGCC 5271  
QY 5276 TTCTACACTGCTGCAAGACCAAGTCCGGAACAGCTACTATACATCGGGCCGATGGCTCC 5335  
Db 5272 TTCTAC-----GACCAAGTCCGGAACAGCTACTATACATCGGGCCGATGGCTCC 5319  
QY 5336 TTGCGGCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACCTTCTG 5395  
Db 5320 TTGCGGCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACCTTCTG 5379  
QY 5396 GCTGGCACCGTCAACCCCAACCGTGGCAAGAGGAATGTACCGTCCCCATCGACAAACCGC 5455  
Db 5380 GCTGGCACCGTCAACCCCAACCGTGGCAAGAGGAATGTACCGTCCCCATCGACAAACCGC 5439  
QY 5456 CTCAAACCTGTGGAGTGGCGCCAGCGCAAGAGAGAGGTCTGGGCCAGGTCACTGTCTTT 5515  
Db 5440 CTCAAACCTGTGGAGTGGCGCCAGCGCAAGAGAGAGGTCTGGGCCAGGTCACTGTCTTT 5499  
QY 5516 GGGCGCCGGTGGCGGTG-----CACAAACGAAATCTCTATCTCTGGACTTTGAT 5566  
Db 5500 GGGCGCCGGTGGCGGTGCTCCAGGTTCACAAACGAAATCTCTATCTCTGGACTTTGAT 5559  
QY 5567 CGCGTAACACGACAGAGAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTG 5626  
Db 5560 CGCGTAACACGACAGAGAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTG 5619  
QY 5627 TACGACAGGCGGGCGCCAGCCCTCTGCTCACCCAGCAGCAGGTGAATGGTGTCAAC 5686  
Db 5620 TACGACAGGCGGGCGCCAGCCCTCTGCTCACCCAGCAGCAGGTGAATGGTGTCAAC 5679  
QY 5687 GTGACATACCTCCCTGGGGTTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGA 5746  
Db 5680 GTGACATACCTCCCTGGGGTTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGA 5739  
QY 5747 ATGGAATACGACAGGCGGGCGCCAGCATCATCCAGGATCTTCGCTGATGGGAAGACATGG 5806  
Db 5740 ATGGAATACGACAGGCGGGCGCCAGCATCATCCAGGATCTTCGCTGATGGGAAGACATGG 5799  
QY 5807 AGCTACACATCTTAGAGAAAGTCCATGGTGTCTACTACACAGCCAGGAGGAGTATATC 5866  
Db 5800 AGCTACACATCTTAGAGAAAGGAGGTG----- 5827

QY	5867	TTTGAGTTTCGACAAAGATGACCGCCTCTCTTCTGTGACGATGCCAAACGTGGCGGGCAG	5926
Dd	5828	-TCGAGTTCGACAAAGATGACCGCCTCTCTTCTGTGACGATGCCAAACGTGGCGGGCAG	5886
QY	5927	ACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGC	5986
Dd	5887	ACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGC	5946
QY	5987	AATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCACTCTTACACCTTCTACCTG	6046
Dd	5947	AATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCACTCTTACACCTTCTACCTG	6006
QY	6047	GGCACTGGCCGACAGGTGATATACAAGTATGGCAACTGTCAAAGCTGGCAGAGCGTC	6106
Dd	6007	GGCACTGGCCGACAGGTGATATACAAGTATGGCAACTGTCAAAGCTGGCAGAGCGTC	6066
QY	6107	TATGACACCAACCAAGTTCAGTTTCACTATGACGAGCGGACGATGCTGAAGACCATC	6166
Dd	6067	TATGACACCAACCAAGTTCAGTTTCACTATGACGAGCGGACGATGCTGAAGACCATC	6126
QY	6167	AACCTACAGAATGAGGGTTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCTGATT	6226
Dd	6127	AACCTACAGAATGAGGGTTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCTGATT	6186
QY	6227	GACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGCTCAACGCCCGTTTTCACTACAAC	6286
Dd	6187	GACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGCTCAACGCCCGTTTTCACTACAAC	6246
QY	6287	TATGACAAACAGCTTCCGGGTGACCAAGCATGACGCTGTGATCAACGAGACCCCACTGCC	6346
Dd	6247	TATGACAAACAGCTTCCGGGTGACCAAGCATGACGCTGTGATCAACGAGACCCCACTGCC	6306
QY	6347	ATTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGT	6406
Dd	6307	ATTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGT	6366
QY	6407	GTCATTTACTATGACATTAACAGATCATCACCACAGCTGTCTATGACCCACACCAAGCAT	6466
Dd	6367	GTCATTTACTATGACATTAACAGATCATCACCACAGCTGTCTATGACCCACACCAAGCAT	6426
QY	6467	TTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAC	6526
Dd	6427	TTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAC	6486
QY	6527	TGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGGAGTGAAGGTAGGA	6586
Dd	6487	TGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGGAGTGAAGGTAGGA	6546
QY	6587	CCCTACGCCAATACCACTCGCTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACA	6646
Dd	6547	CCCTACGCCAATACCACTCGCTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACA	6606
QY	6647	GTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCAC	6706
Dd	6607	GTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCAC	6666
QY	6707	TTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGC	6766
Dd	6667	TTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGC	6726
QY	6767	ATCACTCGGCTGGGTGAGTGCAATACAAGATGAGTGGCTTCTCTGAGGCAGCGG	6826
Dd	6727	ATCACTCGGCTGGGTGAGTGCAATACAAGATGAGTGGCTTCTCTGAGGCAGCGG	6786
QY	6827	GGCGGTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCCCTACAACCGGGCT	6886
Dd	6787	GGCGGTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCCCTACAACCGGGCT	6846
QY	6887	GGCAGCTGGAGTGTCAAGTACCGCTACGATGGCGGGCGGCGTGTCCAGCAAGAGC	6946
Dd	6847	GGCAGCTGGAGTGTCAAGTACCGCTACGATGGCGGGCGGCGTGTCCAGCAAGAGC	6906
QY	6947	AGCCACAGCCACACCTGCAGTCTTCTATGCAGACCTGCACCAACCCACCAAGGTCAAC	7006
Dd	6907	AGCCACAGCCACACCTGCAGTCTTCTATGCAGACCTGCACCAACCCACCAAGGTCAAC	6966
QY	7007	CACCTGTACAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGAAGGACAC	7066
Dd	6967	CACCTGTACAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGAAGGACAC	7026
QY	7067	CTCTTTGCCATGGAGCTGAGAGTGGTGTGATGAGTTTACATAGCTTGTGACAACATCGGG	7126
Dd	7027	CTCTTTGCCATGGAGCTGAGAGTGGTGTGATGAGTTTACATAGCTTGTGACAACATCGGG	7086
QY	7127	ACCCCTCTTGTCTCTTTTAGTGAACAGGTTTGTGATGATCAAGCAAAATCCTGTACACAGCC	7186
Dd	7087	ACCCCTCTTGTCTCTTTTAGTGAACAGGTTTGTGATGATCAAGCAAAATCCTGTACACAGCC	7146
QY	7187	TATGGGAGATCTACATGGATACCAACCCCAACTTTTTCAGATCATATAGGCTACCATGGT	7246
Dd	7147	TATGGGAGATCTACATGGATACCAACCCCAACTTTTTCAGATCATATAGGCTACCATGGT	7206
QY	7247	GGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGGCGAGATTATGATGTGCTG	7306
Dd	7207	GGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGGCGAGATTATGATGTGCTG	7266
QY	7307	GCCGACGCTGACCTAGCCACAGACCAAGCTGTGGAAGCACCTTAGTAGCAGCAACGTC	7366
Dd	7267	GCCGACGCTGACCTAGCCACAGACCAAGCTGTGGAAGCACCTTAGTAGCAGCAACGTC	7326
QY	7367	ATGCTTTTAATCTCTATATGTTCAAAAAACAACCCCATCAGCAACTCCCAGGACATC	7426
Dd	7327	ATGCTTTTAATCTCTATATGTTCAAAAAACAACCCCATCAGCAACTCCCAGGACATC	7386
QY	7427	AAGTGCTTCATGACAGATGTTAAACAGCTGGCTGCCTTTCAGATTCCAGCTACACAAC	7486
Dd	7387	AAGTGCTTCATGACAGATGTTAAACAGCTGGCTGCCTTTCAGATTCCAGCTACACAAC	7446
QY	7487	GTGATCCCTGGTTATCCCAAACAGACATGGATGCCATGGAACCCCTCCTACGAGCTCATC	7546
Dd	7447	GTGATCCCTGGTTATCCCAAACAGACATGGATGCCATGGAACCCCTCCTACGAGCTCATC	7506
QY	7547	CACACACAGATGAAAAACGACGAGTGGGCAACACAGCA-----	7583
Dd	7507	CACACACAGATGAAAAACGACGAGTGGGCAACACAGCAAGGTAATTCCTGCACAAGGTGC	7566
QY	7584	-AGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAACAGCTCAAGGCCCTTTGTCACTTA	7642
Dd	7567	CAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAACAGCTCAAGGCCCTTTGTCACTTA	7626
QY	7643	GAACGGTTTGACCAAGCTCTATGGCTCCACAATCACAGCTGCCAGAGGCTCCAAAGACC	7702
Dd	7627	GAACGGTTTGACCAAGCTCTATGGCTCCACAATCACAGCTGCCAGAGGCTCCAAAGACC	7686
QY	7703	AAGAAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGAT	7762
Dd	7687	AAGAAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGAT	7746
QY	7763	GGCCGAGTGACCACAGACATCATCAGTGTGGCCAAATGAGGATGGGCGAAGGGTTGCTGCC	7822
Dd	7747	GGCCGAGTGACCACAGACATCATCAGTGTGGCCAAATGAGGATGGGCGAAGGGTTGCTGCC	7806
QY	7823	ATCTTGAACCATGCCCACTACCTAGAGAACCTGCACCTTCAACATTGATGGGGTGATACC	7882
Dd	7807	ATCTTGAACCATGCCCACTACCTAGAGAACCTGCACCTTCAACATTGATGGGGTGATACC	7866
QY	7883	CATTACTTTGTGAAACCAAGGACCTTCAGAAGGTGACTGGCCATCCTGGGCTCAGTGGG	7942
Dd	7867	CATTACTTTGTGAAACCAAGGACCTTCAGAAGGTGACTGGCCATCCTGGGCTCAGTGGG	7926
QY	7943	GGCGGCGGAACCCCTGGAGAATGGGGTCAACGTCACCTGTGTCCAGATCAACACAGTACTT	8002
Dd	7927	GGCGGCGGAACCCCTGGAGAATGGGGTCAACGTCACCTGTGTCCAGATCAACACAGTACTT	7986
QY	8003	AATGCAGGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAAC	8062

D <b>b</b>	7987	AATGGCAGGACTAGACGCTACACAGACATCCAGTCCAGTACGGGCAC	8046
QY	8063	ACACGCTACGGGACAACGTTCGGATGAGGAGAAGGCACCGGTCTCTGGAGCTGCCCGGCAG	8122
D <b>b</b>	8047	ACACGCTACGGGACAACGTTCGGATGAGGAGAAGGCACCGGTCTCTGGAGCTGCCCGGCAG	8106
QY	8123	AGAGCCGTGCGCCAAAGCTGGGCCCGCGAGCACACAGACTGCCGGAAGGGGAGGAAGGC	8182
D <b>b</b>	8107	AGAGCCGTGCGCCAAAGCTGGGCCCGCGAGCACACAGACTGCCGGAAGGGGAGGAAGGC	8166
QY	8183	CTGCGGGCCTGGACAGAGGGGGAGAAAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAAGGC	8242
D <b>b</b>	8167	CTGCGGGCCTGGACAGAGGGGGAGAAAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAAGGC	8226
QY	8243	TACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCAGAACTGTACAGACAGCGCCAAC	8302
D <b>b</b>	8227	TACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCAGAACTGTACAGACAGCGCCAAC	8286
QY	8303	AACATCCACTTCATGAGACAGACGCGAGATGGGCCGGAGGTGACAGAGGAC	8354
D <b>b</b>	8287	AACATCCACTTCATGAGACAGACGCGAGATGGGCCGGAGGTGACAGAGGAC	8338

## RESULT 4

```

US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18

```

```

; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Query Match          93.1%; Score 7779.2; DB 13; Length 8438;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 8111; Conservative 0; Mismatches 203; Indels 100; Gaps 11;

QY 32 GCCATGGACGTGAAGGAGAGGAAGCCCTTACCGCTCGCTGACCCCGCGCCGCGACGCCGAG 91
Dbb 1 GCCATGGACGTGAAGGAGAGGAAGCCCTTACCGCTCGCTGACCCCGCGCCGCGACGCCGAG 60

QY 92 CGCCGCTACACCAAGCTCGTCCGCGGACAGCGAGGAGGGCAAGCCCGCAGAAATCGTAC 151
Dbb 61 CGCCGCTACACCAAGCTCGTCCGCGGACAGCGAGGAGGGCAAGCCCGCAGAAATCGTAC 120

QY 152 AGTCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCCTATGGCAGCCGC 211
Dbb 121 AGTCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCCTATGGCAGCCGC 180

QY 212 GTC AAGGACATTGTGCGCGCAGGAGGCCGAGGAATTTGCGGCACAGGTGCCAATTTCACC 271
Dbb 181 GTC AAGGACATTGTGCGCGCAGGAGGCCGAGGAATTTGCGGCACAGGTGCCAATTTCACC 240

QY 272 CTGCGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTACGGGACCCTGTACCGGCACAGAC 331
Dbb 241 CTGCGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTACGGGACCCTGTACCGGCACAGAC 300

QY 332 ATTGGCCTGCCCAATGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCT 391
Dbb 301 ATTGGCCTCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCT 360

QY 392 GACACGGTGCTGTCCCTGAGCA CCCCGTGGCTCTGTGGGGCCGGAGCACACGGTCAAGG 451
Dbb 361 GACACGGTGCTGTCCCTGAGCA CCCCGTGGCTCTGTGGGGCCGGAGCACACGGTCAAGG 420

QY 452 CGCAGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACTCACCGACACCGAGCAT 511
Dbb 421 CGCAGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACTCACCGACACCGAGCAT 480

QY 512 GAAACACTGAGACTGATCATCCGGCGGCCCTGCAGAACCAACGCGGCTCCGACGCCG 571
Dbb 481 GAAACACTGAGACTGATCATCCGGCGGCCCTGCAGAACCAACGCGGCTCCGACGCCG 540

QY 572 CGCCCGCGCTCTCGACCGCCACACCCCCCAACCCAGCACCAACGCGGCTCCATTAACTCC 631
Dbb 541 CGCCCGCGCTCTCGACCGCCACACCCCCCAACCCAGCACCAACGCGGCTCCATTAACTCC 600

QY 632 CTGAACCGGGGCAACTTTCAGCCGAGGAGCAACCCAGACCCCGCCCGCCACCGGACCACTCG 691
Dbb 601 CTGAACCGGGGCAACTTTCAGCCGAGGAGCAACCCAGACCCCGCCCGCCACCGGACCACTCG 660

QY 692 CTCTCCGGAGAGCCCCCTGCCCCGGCGCCAGGAGCCTGCCCCACGCCCCAGGAGAACTGG 751
Dbb 661 CTCTCCGGAGAGCCCCCTGCCCCGGCGCCAGGAGCCTGCCCCACGCCCCAGGAGAACTGG 720

QY 752 CTGCTCAACAGCAACATCCCCCTGGAGACCAGGAACCTAGGCAAGCAGCCATTTCCTAGGG 811
Dbb 721 CTGCTCAACAGCAACATCCCCCTGGAGACCAGGAACCTAGGCAAGCAGCCATTTCCTAGGG 780

QY 812 ACATTGCAGGACAACCTCATTTGAGATGGACATTTCTCGGCGCCTCCCGCCATGATGGGGCT 871
Dbb 781 ACATTGCAGGACAACCTCATTTGAGATGGACATTTCTCGGCGCCTCCCGCCATGATGGGGCT 840

QY 872 TACAGTGACGGGACTTCCTCTTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCCACA 931
Dbb 841 TACAGTGACGGGACTTCCTCTTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCCACA 900

QY 932 TCACCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCTCCGCCCGACCCCTGCC 991

```

Db 901 TCACAGGGTACCCACTGACGTCACGACAGTGTACTCTCCTCCGCCCCGACCCCGTGCCTGCCC 960  
QY 992 CGCAGCACCTTCGCCCGCGCGCTTTAACCTCAAGAACCCCTCAAGTACTGTAACTGG 1051  
Db 961 CGCAGCACCTTCGCCCTGGCCCGGCTTTAACCTCAAGAACCCCTCAAGTACTGTAACTGG 1020  
QY 1052 AAGTGGCAGCCCTGAGCGCATCGTCACTCTCAGCCACTCTGGTCACTCCTGCTGGGCATAC 1111  
Db 1021 AAGTGGCAGCCCTGAGCGCATCGTCACTCTCAGCCACTCTGGTCACTCCTGCTGGGCATAC 1080  
QY 1112 TTTGTGGCCATGCACCTGTTTGGCCTAAACTGGCACTGCGACCCGATGGAGGGCAGATG 1171  
Db 1081 TTTGTGGCCATGCACCTGTTTGGCCTAAACTGGCACTGCGACCCGATGGAGGGCAGATG 1140  
QY 1172 -----TATGAGATCACGGAGGACACAGCCAGCACTTGGCCTGTGCCAACCGACGTCTC 1224  
Db 1141 ACGGATTTATGAGATCACGGAGGACACAGCCAGCACTTGGCCTGTGCCAACCGACGTCTC 1200  
QY 1225 CCTATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGAAAGGCAAGGAACCCAC 1284  
Db 1201 CCTATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGAAAGGCAAGGAACCCAC 1260  
QY 1285 AGAAGGAAAGCCAGTAGTCTTTTCCAGAGGACAGTTTCATAGATTCTGGAGAAATTGA 1344  
Db 1261 AGAAGGAAAGCCAGTAGTCTTTTCCAGAGGCAAGTTTCATAGATTCTGGAGAAATTGA 1320  
QY 1345 TGTGGGAAGGCGAGCCTCCAGAAGATTCCCTCTGGCACTTTCTGGAGATCTCAAGTGTT 1404  
Db 1321 TGTGGGAAGGCGAGCTTCCAGAAGATTCCCTCTGGCACTTTCTGGAGATCTCAAGTGTT 1380  
QY 1405 CATAGACCATCCTGTGCACTCTGAAATTCAATGTCTCTCTGGAAAGGAGCCCTGGTTGG 1464  
Db 1381 CATAGACCATCCTGTGCACTCTGAAATTCAATGTCTCTCTGGAAAGGAGCCCTGGTTGG 1440  
QY 1465 CATTATGGCAGAAAGCCCTCCCTCCTTCACATACACAGTTTGACTTTGTGGAGCTGCT 1524  
Db 1441 CATTATGGCAGAAAGCCCTCCCTCCTTCACATACACAGTTTGACTTTGTGGAGCTGCT 1500  
QY 1525 GGATGGCAGGAGGCTCCTAACCCAGGAGGCGCGAGCCCTAGAGGGACCCCGCCAGTC 1584  
Db 1501 GGATGGCAGGAGGCTCCTAACCCAGGAGGCGCGAGCCCTAGAGGGACCCCGCCAGTC 1560  
QY 1585 TCGGGGAACGTGTCGCCCTCCAGCCATGAGCAGGCTTCATCCAGTATTGGATTTCAGG 1644  
Db 1561 TCGGGGAACGTGTCGCCCTCCAGCCATGAGCAGGCTTCATCCAGTATTGGATTTCAGG 1620  
QY 1645 AATCTGGCACTTGGCTTTTACAATGACGGAAGGAGTCAAGAAGTGTTCTCTTCTCAC 1704  
Db 1621 AATCTGGCACTTGGCTTTTACAATGACGGAAGGAGTCAAGAAGTGTTCTCTTCTCAC 1680  
QY 1705 CACTGCCATTGATGCGGTGGATAACTGCCCCAGCAACTGCTATGGCAATGGTGACTGCAT 1764  
Db 1681 CACTGCCATT-----CTTGATTCCCTGGGCTCTCTGTTTGGGTGATGGAAATGCGT 1731  
QY 1765 CTCTGGGACCTGCCACTGCTTCTCTGGGTTTCTCTGGCCCGGACTGTGGCAGAGCCCTCCTG 1824  
Db 1732 TTTCTGGAACCTTGCCATTGTTTCCAGGATTCTGGGTCCGGAATGTTCAAGAGCCGCTG 1791  
QY 1825 CCCCCTGCTCTGTAGCGGAAATGGCCAATACATGAAAGGCAGATGCTTGTGCCACAGTGG 1884  
Db 1792 TCCAGTGTATTGTAGTGGCAACGGGCGAGTACTCCAAGGGCCGCTGCTGTGTTTCAGCGG 1851  
QY 1885 CTGGAAGGGCGTGTAGTGCATGTGCCCCACCAACCACTGTATCGATGTGGCCTGCAGCAA 1944  
Db 1852 CTGGAAGGGCACCGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGG 1911  
QY 1945 CCATGGCACCTGCATCACGGGCACCTGCATCTGCAACCTTGGCTACAGGGCGAGAGCTG 2004  
Db 1912 TCGTGGGATTTGTATCATGGGCTCCTGTGCTTGCAACTCAGGATACAAAGGAGAAAGTTG 1971  
QY 2005 TGAGGAAGTGACTGCATGGACCCCACTGTTTCAGCCCGGGTGTCTGCGTGAGAGCGCA 2064

Db 1972 TGAAGAAGCTGACTGTATAGACCCCTGGGTGTTCTAATCATGTGTGTGTATCCACGGGGA 2031  
QY 2065 ATGCCAATTGCTTTTGTGGATGGGAGGCACCAACTGCGAGACCCCCAGGGCCACATGCTT 2124  
Db 2032 ATGTCACTGCAGTCCAGGATGGGAGGTAGCAATTGTGAATACTGAAGACCATGTGTCC 2091  
QY 2125 AGACCACTGTTTCAAGCCACCGAACTTCTTCCCGGACACCGGGCTTTGCAGCTGTGACCC 2184  
Db 2092 AGACCACTGCTCCGGCCACCGAACTATCTTCAAGAAAGTGGCTCTCGCACGTGTGACCC 2151  
QY 2185 AAGCTGGAATGGACACGACTGTCTTATCGAGATCTGTGCTGCCGACTGTGTGGCCATGG 2244  
Db 2152 TAACTGGAATGGCCAGACTGCTCAAAACGAAATATGTTCTGTGGACTGTGGCTCACACGG 2211  
QY 2245 CGTGTGCTAGGGGACCTTGGCGCTGCGAGGATGGCTGGATGGGGCAGCCCTGCGACCA 2304  
Db 2212 CGTTTGCATGGGGGACGCTGTCGCTGTGAAGAAGGCTGGACGGGCCAGCCCTGTAAATCA 2271  
QY 2305 GCGGGCTGCCACCCCGCTGTGCGAGCATGGGACCTGCGCGACGGCAAGTGCAGTG 2364  
Db 2272 GAGAGCTGCCACCCCGCTGTGCGAGCACGGGACCTGCGCGACGGCAAGTGCAGTG 2331  
QY 2365 CAGCCCTGGCTGGAATGGCGAACACTGCACCATC-----GCTCACTATCTGGATAGGGT 2418  
Db 2332 CAGCCCTGGCTGGAATGGCGAACACTGCACCATCTCCTTAGCTCACTATCTGGATAGGGT 2391  
QY 2419 AGTTAA-----AGAGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGA 2472  
Db 2392 AGTTAAACTTTTCAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGA 2451  
QY 2473 CCTGAATGGTTGGCACTGCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTC 2532  
Db 2452 CCTGAATGGTTGGCACTGCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTC 2511  
QY 2533 CATGGAGACTGCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTGGTGAAGTGCAT 2592  
Db 2512 CATGGAGACTGCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTGGTGAAGTGCAT 2571  
QY 2593 GGACCCTGACTGCTGCCAGCCCTGTGCCATATCAACCCGCTGTGCCCTGGCTCCCC 2652  
Db 2572 GGACCCTGACTGCTGCCAGCCCTGTGCCATATCAACCCGCTGTGCCCTGGCTCCCC 2631  
QY 2653 TAACCCCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTACAGCAGAACCTACACTC 2712  
Db 2632 TAACCCCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTACAGCAGAACCTACACTC 2691  
QY 2713 CTCTATGACCCGATCAAGTTCTCTGTTGGCAGGACAGCACACATAATCCCCGGGA 2772  
Db 2692 CTCTATGACCCGATCAAGTTCTCTGTTGGCAGGACAGCACACATAATCCCCGGGA 2751  
QY 2773 GAAACCCCTTTGATGGAGGCACTGTTGTGTTATTCGTGGCCAAAGTGTGATGACATCAGATGG 2832  
Db 2752 GAAACCCCTTTGATGGAGGCACTGTTGTGTTATTCGTGGCCAAAGTGTGATGACATCAGATGG 2811  
QY 2833 AACCCTCTGTTGGTGAACATCAGTTTTTGTCAATAACCCCTCTCTTTTGGATATACAAT 2892  
Db 2812 AACCCTCTGTTGGTGAACATCAGTTTTTGTCAATAACCCCTCTCTTTTGGATATACAAT 2871  
QY 2893 CAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCCTGCG 2952  
Db 2872 CAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATCCTGCG 2931  
QY 2953 GTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTT 3012  
Db 2932 GTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCCCTGTGGCTGCCATGGGATCGCTT 2991  
QY 3013 CTTTGTTCATGGAAACCATCATCATGAGACATGAGGAGAAATGAGATTCCAGCTGTGACCT 3072  
Db 2992 CTTTGTTCATGGAAACCATCATCATGAGACATGAGGAGAAATGAGATTCCAGCTGTGACCT 3051  
QY 3073 GAGCAATTTTGGCGGCCCAACCCAGTCTCTTCCATCCCCACTGAGCTCCTTCGCCAG 3132  
Db 3052 GAGCAATTTTGGCGGCCCAACCCAGTCTCTTCCATCCCCACTGAGCTCCTTCGCCAG 3111

QY	3133	CTCCTGTGCAGAGAAAGCCCCCATTTGTCCGGAAATTTCAGGCTTTTGCAGGAGGAAATCTC	3192
Db	3112	CTCCTGTGCAGAGAAAGCCCCCATTTGTCCGGAAATTTCAGGCTTTTGCAGGAGGAAATCTC	3171
QY	3193	TATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCCTGCCTACAAATC	3252
Db	3172	TATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCCTGCCTACAAATC	3231
QY	3253	TGTCCTGAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGGAAGGTGCACCT	3312
Db	3232	TGTCCTGAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGGAAGGTGCACCT	3291
QY	3313	CATGGTAGCGGTGAGGGCCGCTCTTTCAGGAAGTGGTTCCGTGCAGCCCCCAGACCTGTC	3372
Db	3292	CATGGTAGCGGTGAGGGCCGCTCTTTCAGGAAGTGGTTCCGTGCAGCCCCCAGACCTGTC	3351
QY	3373	CTATTATTTCATTTGGACAAGACAGACGCTCTACAAACCAGAAGGTGTTTGGGCTTTTCAGA	3432
Db	3352	CTATTATTTCATTTGGACAAGACAGACGCTCTACAAACCAGAAGGTGTTTGGGCTTTTCAGA	3411
QY	3433	AGCCTTTGTTTCCGTGGTTATGAATATGAATCCTTGCCCCCAGATCTAATCCTGTGGGAAAA	3492
Db	3412	AGCCTTTGTTTCCGTGGTTATGAATATGAATCCTTGCCCCCAGATCTAATCCTGTGGGAAAA	3471
QY	3493	AAGAACAACAGTGCTGCAGGGCTATGAAATTGACGGCTCCAAGCTTGGAGATGGAGCCT	3552
Db	3472	AAGAACAACAGTGCTGCAGGGCTATGAAATTGACGGCTCCAAGCTTGGAGATGGAGCCT	3531
QY	3553	AGACAAACATCATGCCCTCAACATTCAAAGTGGT--ATCCTGCACAAAGGAATGGGA	3609
Db	3532	AGACAAACATCATGCCCTCAACATTCAAAGTGGTGGCATCTCTGCACAAAGGAATGGGA	3591
QY	3610	GAACCAGTTTGTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATCATGGGCAATGGCGCCG	3669
Db	3592	GAACCAGTTTGTGTCTCAGCAGCCTCCTGTCTATTGGGAGCATCATGGGCAATGGCGCCG	3651
QY	3670	GAGAAAGCATCTCCTGCCCCAGCTGCAACGGCCTTGTCTACGGCAACAAGCTCCTGGCCCC	3729
Db	3652	GAGAAAGCATCTCCTGCCCCAGCTGCAACGGCCTTGTCTACGGCAACAAGCTCCTGGCCCC	3711
QY	3730	AGTGGCCCTCACCTGTGGCTCTGACGGAGCCTCTATGTGGGTGATTTCAACTACATTAG	3789
Db	3712	AGTGGCCCTCACCTGTGGCTCTGACGGAGCCTCTATGTGGGTGATTTCAACTACATTAG	3771
QY	3790	AAGGATCTTCCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGG-----AATAAAGA	3843
Db	3772	AAGGATCTTCCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGTTACAGAAATAAGA	3831
QY	3844	TTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGG	3903
Db	3832	TTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGG	3891
QY	3904	GGCGTCTTCTCCTTTCTGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGFGT	3963
Db	3892	GGCGTCTTCTCCTTTCTGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGFGT	3951
QY	3964	GGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCACTGCTCC	4023
Db	3952	GGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCACTGCTCC	4011
QY	4024	CTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCC	4083
Db	4012	CTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCC	4071
QY	4084	CAGGGT-----ATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCAC	4134
Db	4072	CAGGGTCCCCCAGGCATTAAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCAC	4131
QY	4135	CATGATCAGACGCATCGATCAGAAATGGGATCATCTCCACCTCTCGGCTCTAATGATCT	4194
Db	4132	CATGATCAGACGCATCGATCAGAAATGGGATCATCTCCACCTCTCGGCTCTAATGATCT	4191

QY	4195	CACATCAGCCCGGCCCACTCAGCTGTGAAATCTGTCAATGGATATTTCCAGGTAAGA-----	4249
DB	4192		
QY	4250	----CTGAGTGGCCCAAGACTTAGCCATCAACCCCAATGGACAACTCACTTTATGTCCT	4305
DB	4252		
QY	4306	CGACAACAATGTGGTCCTGCAAACTCTGTAAACACCAAGGTGGCATTTGTCGCCGGGAG	4365
DB	4312		
QY	4366	GCCCATGCACTGCCAGGTCCTGSCAATTGACCACTTCCTGCTAAGCAAGGTGGCCATCCA	4425
DB	4372		
QY	4426	CGCAACCCCTGGAGTCAGCCACCGCTTTGGCTGTTTCACACAATGSGGTCTGTATATTGC	4485
DB	4432		
QY	4486	TGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCACCACTAGTGGAGAGATCTC	4545
DB	4492		
QY	4546	ACTCGTTGCTGGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTT	4605
DB	4552		
QY	4606	TTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGT	4665
DB	4612		
QY	4666	GTGTGCTGATGGGGAGCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCG	4725
DB	4672		
QY	4726	GAAGAACAAAGCCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACCA	4785
DB	4732		
QY	4786	GGAGCTCTATCTGTTTGATACCAACCGGAAGCACCTGTGTACACCAAGCCTGCCACAGG	4845
DB	4792	GGAGCTCTATCTGTTTGATACCAACCGGCAAGCACCTGTACACCAAGCCTGCCACAGG	4851
QY	4846	AGACTACCTGTACAACTTCACCTACACTGGGACGGCGACATCACACTCATCAGACAA	4905
DB	4852		
QY	4906	CAATGGCAACATGTTGTTTGTACCTACACTGGGACGGCGACATCACACTCATCAGACAA	4911
DB	4912		
QY	4966	CCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGTGTGAC	5025
DB	4972		
QY	5026	CACACAAGGACACAGTTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAAC	5085
DB	5032		
QY	5086	CAAAAGCAATGAAAAACGGATGGACAACATTTATGAGTACGACAGCTTTGGCCGCTGAC	5145
DB	5092		
QY	5146	AAATGTGACCTTCCCTACTGGCCAGGTGAGCGAGTTTCCGAAGTGATACAGACAGTTCAGT	5205
DB	5152		
QY	5206	GCATGTCCAGGTAGACACTCCAGCAAGGATGATGTACCATTAACCAACCTGTCTGC	5265
DB	5212		
QY	5266	CTCAGGGCCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGC	5325

Db 5272 CTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGC 5331  
QY 5326 CGATGGCTCCTTGGCGCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAGACTGAGCC 5385  
Db 5332 CGATGGCTCCTTGGCGCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAGACTGAGCC 5391  
QY 5386 CCACCTTGTGGCTGGCACCCTGTCACCCCAACCGTGGGCAAGAGGAATGTACGGCTGCCCAT 5445  
Db 5392 CCACCTTGTGGCTGGCACCCTGTCACCCCAACCGTGGGCAAGAGGAATGTACGGCTGCCCAT 5451  
QY 5446 CGACAAACGGCCTCAACCTGGTGGAGTGGCGGCAGCGCAAGAGAGCGCTCGGGGCCAGGT 5505  
Db 5452 CGACAAACGGCCTCAACCTGGTGGAGTGGCGGCAGCGCAAGAGAGCGCTCGGGGCCAGGT 5511  
QY 5506 CACTGTCTTTGGGCGCGGCTGCGGGTG-----CACAAACCGAAATCTCCTATCTCT 5556  
Db 5512 CACTGTCTTTGGGCGCGGCTGCGGGTGCTCCAGGTTCAACACCGAAATCTCCTATCTCT 5571  
QY 5557 GGACTTTGATCGGTAAACACGCACAGAGAAGATCTATGATGACCCACCGCAAGTTCAACCT 5616  
Db 5572 GGACTTTGATCGGTAAACACGCACAGAGAAGATCTATGATGACCCACCGCAAGTTCAACCT 5631  
QY 5617 TCGGATTCTGTACGACAGCGCGGGCGGCCAGCCTCTGGTCACCCAGCAGCAGGCTGAA 5676  
Db 5632 TCGGATTCTGTACGACAGCGCGGGCGGCCAGCCTCTGGTCACCCAGCAGCAGGCTGAA 5691  
QY 5677 TGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCAT 5736  
Db 5692 TGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCAT 5751  
QY 5737 GTCTGAAAGAAATGGAATACGACAGCGCGGCGGCATCACATCCAGGATCTTCGCTGATGG 5796  
Db 5752 GTCTGAAAGAAATGGAATACGACAGCGCGGCGGCATCACATCCAGGATCTTCGCTGATGG 5811  
QY 5797 GAAGACATGGAGCTACACATACTTAGAGA-----AGTCCATGGTGTGCTACT 5844  
Db 5812 GAAGACATGGAGCTACACATACTTAGAGAAGCGAGGTGTCCAGTCCATGGTGTGCTACT 5871  
QY 5845 ACACAGCCAGGAGTATATCTTTGAGTTGACAAGAAATGACCGCCTCTCTTCTGTGAC 5904  
Db 5872 ACACAGCCAGGAGTATATCTTTGAGTTGACAAGAAATGACCGCCTCTCTTCTGTGAC 5931  
QY 5905 GATGCCCAACGTGGCGGCGCAGACACTAGACACCATCCGCTCAGTGGGCTACTACAGAAA 5964  
Db 5932 GATGCCCAACGTGGCGGCGCAGACACTAGACACCATCCGCTCAGTGGGCTACTACAGAAA 5991  
QY 5965 CATCTATCAGCCCTGAGGGCAATGCCCTCAGTCATACAGGACTTCACTGAGGATGGGCA 6024  
Db 5992 CATCTATCAGCCCTGAGGGCAATGCCCTCAGTCATACAGGACTTCACTGAGGATGGGCA 6051  
QY 6025 CCTCCTTACACCTTCTACCTGGGCACITGGCGCAGGGTGATATAAAGTATGGCAAAC 6084  
Db 6052 CCTCCTTACACCTTCTACCTGGGCACITGGCGCAGGGTGATATAAAGTATGGCAAAC 6111  
QY 6085 GTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAAGTTTCACTTATGACGAGAC 6144  
Db 6112 GTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAAGTTTCACTTATGACGAGAC 6171  
QY 6145 GGCAGGCATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACTGACCATCCGCTA 6204  
Db 6172 GGCAGGCATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACTGACCATCCGCTA 6231  
QY 6205 CCGTCAGATTGGGCCCTGATTGACCCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGT 6264  
Db 6232 CCGTCAGATTGGGCCCTGATTGACCCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGT 6291  
QY 6265 CAACGCCCGTTTTTGAATACTATGACACAGCTTCCGGGTGACCAAGCATGCAGGCTGT 6324  
Db 6292 CAACGCCCGTTTTTGAATACTATGACACAGCTTCCGGGTGACCAAGCATGCAGGCTGT 6351  
QY 6325 GATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAAGAC 6384

Db 6352 GATCAACGAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGTGTGAGCAAGAC 6411  
QY 6385 AGAGCAGTTTGGGAAGTTTGGTGTCTATTTACTATGACATTAACCAAGATCATCACACAGC 6444  
Db 6412 AGAGCAGTTTGGGAAGTTTGGTGTCTATTTACTATGACATTAACCAAGATCATCACACAGC 6471  
QY 6445 TGTCTATGACCCACACCAAGCATTTTGTATGATATGCGCAGGATGAAGGAAGTGCACTATGA 6504  
Db 6472 TGTCTATGACCCACACCAAGCATTTTGTATGATATGCGCAGGATGAAGGAAGTGCACTATGA 6531  
QY 6505 GATCTTCCGCTCGCTCATGTACTGTGATGACCGTCCAGTATGATAACATGGGGCGAGTAGT 6564  
Db 6532 GATCTTCCGCTCGCTCATGTACTGTGATGACCGTCCAGTATGATAACATGGGGCGAGTAGT 6591  
QY 6565 GAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGA 6624  
Db 6592 GAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGA 6651  
QY 6625 TGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGCGCGCTACAGCTA 6684  
Db 6652 TGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGCGCGCTACAGCTA 6711  
QY 6685 CGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCAAGGCTCACACCACT 6744  
Db 6712 CGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCAAGGCTCACACCACT 6771  
QY 6745 ACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGA 6804  
Db 6772 ACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGA 6831  
QY 6805 GGATGGCTTCTTGAGGAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCCTGCT 6864  
Db 6832 GGATGGCTTCTTGAGGAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCCTGCT 6891  
QY 6865 CATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCCTGGG 6924  
Db 6892 CATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCCTGGG 6951  
QY 6925 GCGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCACTGCAGTCTTCTATGACAGACCT 6984  
Db 6952 GCGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCACTGCAGTCTTCTATGACAGACCT 7011  
QY 6985 GACCAACCCCAACCAAGTCAACCACTGTGTACAAACCACTCCAGCTCTGAGATCACCTCCCT 7044  
Db 7012 GACCAACCCCAACCAAGTCAACCACTGTGTACAAACCACTCCAGCTCTGAGATCACCTCCCT 7071  
QY 7045 CTACTACGACTTTCRAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGATGAGTTT 7104  
Db 7072 CTACTACGACTTTCRAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGATGAGTTT 7131  
QY 7105 CATAGCTTGTGACAAACATCGGGACCCCTCTTGTGTCTTTTAGTGGAAACAGGTTTATGAT 7164  
Db 7132 CATAGCTTGTGACAAACATCGGGACCCCTCTTGTGTCTTTTAGTGGAAACAGGTTTATGAT 7191  
QY 7165 CAAGCAAAATCCTGTACACAGCCCTATGGGGAGATCTACATGGATACCAACCCCACTTTCA 7224  
Db 7192 CAAGCAAAATCCTGTACACAGCCCTATGGGGAGATCTACATGGATACCAACCCCACTTTCA 7251  
QY 7225 GATCATCATAGGCTACCATGGTGGCCTCTATGATCCCACTCAACCAAGCTTGTCCACATGGG 7284  
Db 7252 GATCATCATAGGCTACCATGGTGGCCTCTATGATCCCACTCAACCAAGCTTGTCCACATGGG 7311  
QY 7285 CCGGCGAGATTATGATGTGCTGGCCGACGCTGGACTAGCCAGACCAAGCTGTGGAA 7344  
Db 7312 CCGGCGAGATTATGATGTGCTGGCCGACGCTGGACTAGCCAGACCAAGCTGTGGAA 7371  
QY 7345 GCACCTTAGTAGCAACAGCTCATGCCCTTTTAAATCTCTATATGTTCAAAAACAACCAACCC 7404  
Db 7372 GCACCTTAGTAGCAACAGCTCATGCCCTTTTAAATCTCTATATGTTCAAAAACAACCAACCC 7431  
QY 7405 CATCAGCAACTCCAGGACATCAAGTGTCTTATGACAGATGTTAAACAGCTGGCTGCAC 7464  
Db 7432 CATCAGCAACTCCAGGACATCAAGTGTCTTATGACAGATGTTAAACAGCTGGCTGCAC 7491





Db 2338 AGCCAGGGCTGGAATGGAGAGCACTGCACTATCGCTACCTATTGGATAAGATAGTTAAA 2397  
Qy 2422 -----TAAAGAGGGTTGGCTGGTGTGCAATGGCAACGGCAGATGTACC 2467  
Db 2398 GACAAGATAGGATATAAAGAGGGTTGCTGCTGTGCAACAGCAATGGAAGATGTACC 2457  
Qy 2468 TTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGAC 2527  
Db 2458 CTGGACCAAAATGGCGGACATTGTGTGTGCCAGCCTGGATGGAGAGGAGCAGGCTGTGAC 2517  
Qy 2528 ACTTCCATGGAGACTGCCCTGGGTGACAGCAAAAGACAATGATGGAGATGGCCTGGTGGAC 2587  
Db 2518 GTAGCCATGGAGACTCTTTGCACAGATAGCAAGGACAATGAAGGGATGGACTCATGTAC 2577  
Qy 2588 TGCATGGACCTGACTGCTGCCCTCCAGCCCCCTGTGCCATATCAACCCGCTGTGCCCTTGGC 2647  
Db 2578 TGCATGGATCCCGATTGCTGCCCTACAGAGTTCTTGCCAGAAATCAGCCCTATTGTGCGGGA 2637  
Qy 2648 TCCCCTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCTGTGTACAGCAGAAACCTA 2707  
Db 2638 CTGCCGGATCCTCAGGACATCATTTAGCCAAAAGCCTTCAATCGCTTCTCAGCAAGCTGCC 2697  
Qy 2708 CACTCCTTCTATGACCGCATCAAGTTCTCTGTTGGCAGGAGCAGCACGACATATCCCC 2767  
Db 2698 AAATCCTTTTATGATCGAATCAGTTTCTTATAGGATCTGATAGCACCCATGTTATACCT 2757  
Qy 2768 GGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGCCAAAGTGATGACATCA 2827  
Db 2758 GGAGAAAGTCTTCAATAAGAGCCTTGATCTGTCTATCAGAGGCCAAAGTACTGACTGCT 2817  
Qy 2828 GATGGAACCCCTGGTGGTGTGAACATCAGTTTGTCAATAAACCCCTCTCTTTGGATAT 2887  
Db 2818 GATGGAACCTCACTTATTGGAGTAAATGCTCGTTTTTCCATTACCCAGAAATATGGATAT 2877  
Qy 2888 ACAATCAGCAGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCATCATC 2947  
Db 2878 ACTATTACCCGCGAGGACGGAATGTTGACTTGGTGCAAAATGGTGGGCCCTCTCTAACT 2937  
Qy 2948 CTGCGGTTGAGCGGGCACCTTTTCATCACACAGGAGCACACCTGTGGCTGCCATGGAT 3007  
Db 2938 TTGGTATTTGACGATCCCCATTCCTCACTCAGTATCATACTGTGTGGATTCCATGGAAT 2997  
Qy 3008 CGCTTCTTTTGTATGAAACCATCATATGAGACATGAGGAGAAATGAGATTCCCAGTGT 3067  
Db 2998 GTCTTTTATGTATGGATACCCCTAGTCATGGAGAAAGAGAAATGACATTCCCAGTGT 3057  
Qy 3068 GACCTGAGCAATTTTGGCCGCCCAACCCAGTCTCTCTCCATCCCCACTGACGTCCTTC 3127  
Db 3058 GATCTGAGTGGATTCGTGAGGCCAAATCCCATCATTTGTGTCACTACCTTTATCCACCTT 3117  
Qy 3128 GCCAGCTCCTGTGAGAGAAAGGCCCAATTGTGCCGGAATTCAGGCTTTTGACAGGAGAA 3187  
Db 3118 TTCAGATCTTCTCCTGAAGACAGTCCCATCATTCCTCCGAAACACAGGTACTCCACGAGAA 3177  
Qy 3188 ATCTCTATCTGTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCCTGGCTAC 3247  
Db 3178 ACTACAAATTCAGGAAACAGATTTGAAACTCTCTACTTGTAGTTCAGAGCTGCAGGAT 3237  
Qy 3248 AAATCTGTCTGAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTG 3307  
Db 3238 AAGTCAGTTCTCAAGATCACCATGACCCAGTCTATTATTCCATTAAATTTAATGAAGTT 3297  
Qy 3308 CACCTCATGTGAGCGGTGGAGGCCGCTCTTTCAGGAAAGTGGTTCGCTGACGCCAGAC 3367  
Db 3298 CATCTTATGGTAGCTGTAGTAGGAAGACTCTTCCAAAAGTGGTTTCTGCCTCACCAAAAC 3357  
Qy 3368 CTGTCTTATTTTTCATTTGGGACACAGACAGCTCTACAAACAGAAAGGTGTTTGGGCTT 3427  
Db 3358 TTGGCCCTATACCTTTCATATGGGATAAAACAGATGCATATAATATCAGAAAAGTCTATGGTCTA 3417  
Qy 3428 TCAGAAGCCTTTGTTTCCGTGGTTATGAATATGAATCCTGCCAGATCTAATCCTGTGG 3487

Db 3418 TCTGAAGCTGTGTGTCAGTTGGATATGAGTATGAGTCTGTGTTTGGACCTGACTCTGTGG 3477  
Qy 3488 GAAAAAAGAAACACAGTGTGACGGCTATGAAATTGACGGCTCCAAGCTTGGAGGATGG 3547  
Db 3478 GAAAAAGAGACTGCCATTCTGACGGCTATGAAATTGGATGCGTCCAACATGGGTGGCTGG 3537  
Qy 3548 AGCCTAGACAAACATCATGCCCTCAACATTCAAAGTGTATCTCTGCACAAAGGGAATGGG 3607  
Db 3538 ACATTAGATAAAACATCACGTGCTGGATGTACAGAACGGTATACTGTACAAGGGAACGGG 3597  
Qy 3608 GAGAACCAAGTTGTGTCTCAGAGCCTCTGTCTCATTTGGGAGCATCATGGCAATGGGCGC 3667  
Db 3598 GAAAAACCAAGTTCTCTCCAGCAGCCTCCAGTCTGTGAGTAGCATCATGGCAATGGGCGA 3657  
Qy 3668 CGGAGAAAGCATCTCCTGCCCCAGCTGCAACGGCCTTGTGACGGCAACAAAGCTCCTGGCC 3727  
Db 3658 AGCGCAGCATTTCTGCCCCAGTTGCAATGGTCAAGCTGATGGTAAACAAGTTACTGGCC 3717  
Qy 3728 CCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATT 3787  
Db 3718 CCAGTGGCGCTAGCTTGTGGGATCGATGGCAGTCTGTACGTAGGCGATTTCAACTACGTG 3777  
Qy 3788 AGAAGGATCTTCCCCTCTGGAATGTCAACCAACATCCTAGAGCTGAGGAATAAAGATTTT 3847  
Db 3778 CGCGCGATATTCCCTTCTGGAATGTAAACAAGTGTCTTAGAACTAAGAAATAAAGATTTT 3837  
Qy 3848 AGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGTGGGCGC 3907  
Db 3838 AGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACGGGAGAT 3897  
Qy 3908 GTCTTCTTCTTGACAGCAACAGCCGGCGGTCTTTAAATCAAGTCCACTGTGTGTGTG 3967  
Db 3898 CTGTACGTTTCTGACACAAACACCCCGCAGAAATTTATCGCCCAAAAGTCACTTACGGGGCA 3957  
Qy 3968 AAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAAGTGCCTCCCTTT 4027  
Db 3958 AAAGACTTGACTAAATATGCAAGTCTGTCGAGGACAGGGGAGCAATGCCCTCCGTTT 4017  
Qy 4028 GATGACACTCGCTCGGGGATGTTGGGAAGGCCACAGAAAGCCACACTCACCATCCAGG 4087  
Db 4018 GACGAGCGAGATGTGGGATGGAAGGAAGGCCCTGGAAGCCACACTCATGAGTCCCAAA 4077  
Qy 4088 GGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCTGTGATGGCACCATGATCAGACGC 4147  
Db 4078 GGAATGGCAGTTGATAAGAAATGGATTAAATCTACTTTGTTGATGGAACCATGATTAGGAA 4137  
Qy 4148 ATCGATCAGAATGGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATCAGCCCGG 4207  
Db 4138 GTTGACCAAAATGGAATCATATCAACTCTTCTGGCTCTAACGATTTGACTTTCAGCCAGA 4197  
Qy 4208 CCACTCAGCTGTGATTCTGTCAATGATATTTCCAGGTAAGACTGGAGTGGCCCCACAGAC 4267  
Db 4198 CCTTTAACTTGTGACACCAGCATGCAATCAGCCAGGTACGCTGGAATGGCCACTGAC 4257  
Qy 4268 TTAGCCATCAACCCCAATGGACAACCTCACTTTATGTCTCGACAAACAAATGTGTCTGCAA 4327  
Db 4258 CTAGCCATTAAACCTTATGGATAACTCCATTTATGTCTCTGGATAAATAATGTAGTTTACAG 4317  
Qy 4328 ATCTCTGAAAAACCCAGGTGCGCATTTGCGCGGAGGCCCATGCACTGCCAGGTCCCT 4387  
Db 4318 ATCACTGAAAAATCGTCAAGTTCGATTCGCTGTGACGGGCCCATGCACTGTCAAGTTCCC 4377  
Qy 4388 GGCATTGACCACTTCTCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCCACC 4447  
Db 4378 GGAGTGG---AATATCTGTGGGAAGCACGCGGTGCAGACAAACACTGGAATCAGCCACT 4434  
Qy 4448 GCTTTGGCTGTTCACACAATGGGCTCTGTATATTGCTGACACTGATGAGAAAAAGATC 4507  
Db 4435 GCCATTGCTGTCTTACAGTGGGCTCTGTACATTACTGAAACTGATGAGAAGAAATTT 4494  
Qy 4508 AACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCCCCCAGT 4567  
Db 4495 AACCGGATAAGGCAGGTCAACAACAGATGGAGAAATCTCTCTTAGTGGCCGGAATACCTTCA 4554



Db 6715 AGTGGCGTCTGACACCCCTTCGCTATGACCTCGAGACAGAAATCACTCGACTGGTGAT 6774  
QY 6785 GTGCAATACAAGATGGATGAGGATGGCTTCCTGAGGCAGCGGGCGGTGATATCTTTGAG 6844  
Db 6775 GTTCAATATCGGTGGATGAAGATGGTTCTTACGTCAAAGGGGCACGAAATCTTTGAA 6834  
QY 6845 TACAACCTCAGCTGGCCTGCTCATCAAGSCTTACAAACCGGCTGCGAGCTGGAGTGTGAGG 6904  
Db 6835 TATAGCTCCAAGGGCTTCTAACTCGAGTTTACAGTAAAGGCACTGGCTGGACAGTGATC 6894  
QY 6905 TACCGCTACGATGGCCTGGGCGCGGTGTCCAGCAAGAGAGCAGCCACACCTG 6964  
Db 6895 TACCGTTATGACGGCTGGGAAGGCGTGTCTAGCAAAACCACTAGGACAGCACCTG 6954  
QY 6965 CAGTTCTTCTATGACAGACCTGACCAACCCCAAGGTCAACCCACCTGTACAAACCACTCC 7024  
Db 6955 CAGTTTTTTTATGCTGACTTAACCTTATCCCACTAGGATTACTCATGTCTACAAACCACTCG 7014  
QY 7025 AGCTCTGAGATCACTCCCTCTACTACGACTTGCAAGGACACCTCTTTGGCCATGGAGCTG 7084  
Db 7015 AGTTCAGAAATTAACCTCCCTGTATTATGATCTCAAGGACATCTTTTGGCCATGGAAATC 7074  
QY 7085 AGCAGTGGTATGAGTTTACATAGCTTGTGACAAACATCGGAGACCCCTCTTTGCTGTCTTT 7144  
Db 7075 AGCAGTGGGATGAATCTATATTGCTATCGGATAACACAGGACACCACTGGCTGTGTTTC 7134  
QY 7145 AGTGAACACAGTTTATGATCAAGCAATPCCTGTACACAGCCTATGGGAGATCTACATG 7204  
Db 7135 AGTAGCAATGGGCTATGCTGAAACAGATTCAGTACACTGCTATGCGAATCTATTTT 7194  
QY 7205 GATACCAACCCCAACTTTTACATCATCAGATCATCAGGCTACCATGGTGGCCTCTATGATCCACTC 7264  
Db 7195 GACTCTAATATTGACTTTCAACTGGTAATGGATTTCATGGTGGCTGTATGACCCACTC 7254  
QY 7265 ACCAAGCTTGTCCACATGGCGCGGAGATTATGATGTGTCGGCGGACGCTGGACTAGC 7324  
Db 7255 ACCAAATTAATCCACTTTGGAGAAAGAGATTATGACATTTTGGCAGGACGGTGGACAACA 7314  
QY 7325 CCAGACCACGAGTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCTCCTTTTAATCTCTAT 7384  
Db 7315 CCTGACATAGAAATCTGGA--AAAGAAATGGGAAGGACCCAGCTCCTTTTAACTTGATC 7371  
QY 7385 ATGTTCAAAACAAACACCCCATCAGCAACTCCAGGACATCAAGTCTTTCATGACAGAT 7444  
Db 7372 ATGTTTAGGAATAACAACCCCTGCAAGCAAAATCCATGACGTGAAGATTACATCACAGAT 7431  
QY 7445 GTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAAACGTCATCCCTGGTTATCCC 7504  
Db 7432 GTTAACAGCTGGCTGGTGACATTTGGTTTCCATCTGCAACATGCTATTCTGGATTCCCT 7491  
QY 7505 AAACACAGACATGGATGCCATGGAACCCCTCCTACGAGCTCATCCACACAGATGAACACG 7564  
Db 7492 GTTCCCAATTTGAATTTAACAGAACCTTCTTACGA-----ACTTGTGAAGAT 7539  
QY 7565 CAGGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAAGCAGCTC 7624  
Db 7540 CAGCAGTGGGATGATATACCGCCCATCTTCGGAGTCCAGCAGCAAGTGGCGGCGCAGGCC 7599  
QY 7625 AAGGCCTTTGTACACCTTAGAACCGTTTGCATCCAGCGGCTCAGTCTTGGCAAGGGGCTC 7684  
Db 7600 AAGGCCTTCTGTCTGGTGGGAAGATGGCCGAGGTGACAGTGGCGCGCGCGCGC 7659  
QY 7685 CAGCAGGCTCCAAAGACCAAGAAGTTTGCATCCAGCGGCTCAGTCTTGGCAAGGGGCTC 7744  
Db 7660 ---GGCGGCGAGTCTGTGGTGTGGTTTCGCCACGGTCAAGTCCGTGATCGGAAGGGCGTC 7716  
QY 7745 AAGTTTGCCTTGAAGGATGGCCGAGTGACACACAGACATCATCAGTGTGGCCAAATGAGGAT 7804  
Db 7717 ATGCTGGCCGTCAGCCAGGGCGCGGTGACAGACCAACGTCGTCTCAACATCGCCAAACGAGAC 7776  
QY 7805 GGGCGAAGGTTGTGCTGCCATCTTGAACCATGCCCACTACCTAGAGAACCTGCACCTCACC 7864

Db 7777 TGCATCAAGGTGGCGCGGTGCTCTCAACACGCCCTTCTACCTGGAGAACCTGCACCTTACC 7836  
QY 7865 ATTGATGGGTGGATACCCATTATCTTTGTGAAAACAGGACCTTTCAGAAAGGTGACCTGGCC 7924  
Db 7837 ATCGAGGGCAAGACACGCACTACTTTCATCAAGACCAACACGCGGAGAGCAGCCTGGGC 7896  
QY 7925 ATCTGGGCTCAGTGGGGGGGGGAAACCTTGAGAAATGGGGTCAACGTCACTGTGTCC 7984  
Db 7897 ACGTGGGTGACCAAGCGCGCGCAAGCGCTGGAGAACCGCATCAACGTGACGGTGTG 7956  
QY 7985 CAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAGTAC 8044  
Db 7957 CAGTCCACCAAGGTGTAACCGGACGACGCGCAGGTTCCGGACGTGGAGATGCACTTC 8016  
QY 8045 GGGCACCTGTCTTGAACACACAGCTACGCGGCAACGTTGGATGAGGAGAAAGCAGCGGTC 8104  
Db 8017 GCGCGCTGGCGCTGCACGTGCGCTACGCGCATGACCTGGACGAGGAGAAAGCGCGCATC 8076  
QY 8105 CTGAGCTGGCGCGGACAGAGAGCCGTGCGGCTGCGCCAAAGCGTGGCGCGGACGAGACTG 8164  
Db 8077 CTGAGCAGCGCGGACGCGCGCTCGCCGGGCTCGCGGCGGCGGAGCAGCAGCGCGTG 8136  
QY 8165 CGGAAGGGAGGAGGCGCTGCGGCGCTGCGGACAGAGGGGAGAGAGCAGGCTGTGAGC 8224  
Db 8137 CGCAGCGGAGGAGGCGCGCTCTGACGAGGCGGAGGCGGAGAGGCGGAGGCGGAGC 8196  
QY 8225 ACAGGCGGCTGCAAGGCTACGACGCTTTTTCGTGATCTCTGTCGAGCAGTACCCAGAA 8284  
Db 8197 GCGGCAAGTGCAGGCTACGACGCGTACTACGTACTCTCGTGGAGCAGTACCCCGAG 8256  
QY 8285 CTGTCAGACAGCGCCAAACATCCACTTCATGAGACAGAGCGAGATGGCGGAGGTGA 8344  
Db 8257 CTGGCGACAGCGCCAAACATCCAGTTCCTGCGGACAGCGAGATCGGCAGGAGTAA 8316  
QY 8345 C 8345  
Db 8317 C 8317

RESULT 6  
US-10-038-854-35  
; Sequence 35, Application US/10038854  
; Publication No. US20040022781A1  
; GENERAL INFORMATION:  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Li, Li  
; APPLICANT: Wolenc, Adam R  
; APPLICANT: Vernet, Corine  
; APPLICANT: Eisen, Andrew J  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Gorman, Linda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Guo, Xiaojia S  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Erik  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: MacDougall, John R  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-230

; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Query Match 41.6%; Score 3473.8; DB 17; Length 8675;
Best Local Similarity 65.2%; Pred. No. 0;
Matches 5445; Conservative 0; Mismatches 2692; Indels 216; Gaps 14;

Qy 35 ATGGACGTGAAGGAGAGAGCCCTTACCGCTCGCTGACCCCGCGC---CGGACGCCGAG 91
Db 151 ATGGATGTGAAGAACGCGAGCCCTTACTGCTCCCTGACCAAGAGCAGACGAGAAGGAA 210
Qy 92 CGCGGCTACACCAGCTCGTCGCGGACAGCGAGGAGGCGCAAG---CCCGCAGAAATCG 148
Db 211 CGCGGCTACACAAATTCCTCCGCGAGACAATGAGGAGTCCGGGTACCCACACAGAAGTCC 270
Qy 149 TACAGCTCCAGCGAGACCTGAAGCCCTACGACCAGGA---CGCCCGCTAGCCTATGGC 205
Db 271 TACAGTTCAGCGAGACATTGAAAGCTTTTGATCATGATTCCTCGCGCTGCTTTACGGC 330
Qy 206 AGCGCGCTCAAGGACATTGTGCGCGAGGAGCGGAGGAATTCGCGCAGAGGTGCCAAC 265
Db 331 AACAGAGTGAAGGATTTGGTTACAGAGAAGCAGACGAGTTCACCTAGACAGACAGAAAT 390
Qy 266 TTCACCCCTGCGGAGCTGGGCTGGAAGAAGTAAACGCCCCCTCACGGACCTGTACCGG 325
Db 391 TTTACCCCTAAGGCAGTGTAGAGTTTGTGAACCGAGCACTCGAAGGACTGGCATTTGT 450
Qy 326 ACAGACATTGGCCTGCCCAATGCGGCTACTCCATGGGGCTGGCTCTGTATGCCGACATG 385
Db 451 GCGGAAATGGGGCTCCCTCACAGAGTTACTCTATCAGTGCAGGTCAGATGCTGTACT 510
Qy 386 GAGGCTGACACGGTGTCTCCCTGAGCACCCCGTGGCTGTGTGGGCGCGGAGCACACGG 445
Db 511 GAAATGAAGCAGTGTATGTCCTCAGAGCATGCCATGAGACTTTGGGGCAGGGGGTCAAA 570
Qy 446 TCAGGGCGCAGCTCCTGCTGTCCAGCGGGCCAAATCCAAATCTCACACTCACCGACACC 505
Db 571 TCAGGCGCAGCTCCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCCTGACAGATACG 630
Qy 506 GAGATGAAAACACTGAGACTGATCATCCGGCGGCTGCAGAACCAACGCGCGGCTCCGG 565
Db 631 GAGCAGGAAAACAAAGTCCGACAGTGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690
Qy 566 ACGCGCGCGCGCTCTCGACGCGCCACACACCCCAACAGCACCCACGCGGCTCCATT 625

Db 691 ACCCTGCAGCCCTTGC CGCCTTCCCATAGCAGCACTCTGCACAGCATCATCCATCATC 750
Qy 626 AACTCCCTGAACCGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCACGAGAC 685
Db 751 ACTTCTCTCAACAGAAACTCCCTGACCAATAGAAGGACAGAGTCCGCGCCCGCGGCT 810
Qy 686 CACTCGCTCTCCGGAGAGCCCTTGC CGCGCGCCAGGAGCCTGCCACGCCCAAGGAG 745
Db 811 GCTTTCGCCCGGAGCTGCAAAAC-----CACACCGAGTCCGTCCAGCTGACGAGAC 861
Qy 746 AACTGGCTGCTCAACAGCAACATCCCTCTGGAGACACAGCAACCTAGGCAAGCAGCATTC 805
Db 862 AGCTGGTCTTGGCAGTATGTACCACTGGAAAGCAGGCATTT----- 905
Qy 806 CTAGGACATTGCAGGACAACTCATTTGAGATGGACATTTCTGGCGCCTCCCGCCATGAT 865
Db 906 ----- 905
Qy 866 GGGCTTACAGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCGGCTCTTCTGTC 925
Db 906 -----CCTATTCAAAACAGGAACAGGTACAAACGCCACTGTTTCAGT 945
Qy 926 ACCACATCACAGGATACCCACTGACGTCCAGCACAGTGTACTCTCTCTCCGCCCGACCC 985
Db 946 ACTGCAACCCAGGATACAAATGGCATCTGGCTCTGTTATTACCACTACTCGGCCA 1005
Qy 986 CTGCCCGCAGCACCTTTCGCCGCGCGCGCTTTAACTCAAGAAAGCCCTCCAAAGTACTGT 1045
Db 1006 CTACCTAGAAACACCCCTATCAAGAAAGTCTTTTAAATTAAGAAAGTCTTCAAAGTACTGT 1065
Qy 1046 AACTGGAAGTGCAGCGCCCTGAGCGCCCATCGTCATCTCAGCCACTCTGGTCACTCTGCTG 1105
Db 1066 AGCTGAAATGCACTGCACTGTGTGCGGTAGGGGTCTCGGTCTCTGGCAATACTCTCTG 1125
Qy 1106 GCATACTTTGTGGCCATGCACCTGTTTGGCCTAAACTGCGACCTGCAGCCGATGGAGGG 1165
Db 1126 TCTTATTTTATAGCAATGCATCTCTTTGGCCTCAACTGCGAGCTACAGCAGACTGAAAT 1185
Qy 1166 CAGATGTATGAGATCACGGAGGACACAGCCAGCAGTGTGGCTGTGCCAACCGACGTCTCC 1225
Db 1186 GACACATTTGAGA-----ATGAAAGTGAATTCGTGATACCATGCCAACAAACACTGTG 1239
Qy 1226 CTATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGAAAGGCAAGAACCAACA 1285
Db 1240 TCATTACCTTCTGGAGACA----- 1258
Qy 1286 GAAGGAAAGCCAGTAGTTTCTTTCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGAT 1345
Db 1259 -ATGGAATAATTAGTGGATTACGCAAGAAATAACACCATAGATTCCGGAGAACTTGAT 1317
Qy 1346 GTGGGAAGCGGAGCCTCCAGAGATTCCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTC 1405
Db 1318 ATTGGCCGAAGAGCAATTCAAGAGATTCTCTCCGGGATCTTCTGGAGATCACAGCTCTTC 1377
Qy 1406 ATAGACCATCTGTGCATCTGAAATTCATATGTCTCTGGGAAAGGACGCCCTGGTTGGC 1465
Db 1378 ATTGATCAGCCACAGTTTCTTAAATTCATATCTCTCTCAGAAAGGATGCATTGATTGA 1437
Qy 1466 ATTTATGGCAGAAAGGCTCCCTCTTCACATA-----CACAGTTTGACTTT 1513
Db 1438 GTATATGGCCGGAAGAGTTACGCGCTTCCCATACTCAGTCTCTCCCGCAGTATGACTTC 1497
Qy 1514 GTGAGCTGTGGATGGCAGGAGGCTCCTAACCCAGGAGGCGGAGCCTTAGAGGGGAC 1573
Db 1498 GTGAGCTCTGGATGGCAGCAGGCTGATTGCCAGAGAGCAGCGGAGCCTGCTTGAGACG 1557
Qy 1574 CCGCGCCAGTCTCGGGAACTGTGCCCGCCCTCCAGCCATGAGACAGGCTTTCATCCAGTAT 1633
Db 1558 GAGAGACCGGGCGGCGAGGAGATCCGTGAGCCTTCATGAGCCGGCTTTATCCAGTAC 1617
Qy 1634 TTGGATTTCAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGAGTCAAGAGTGGTT 1693
Db 1618 TTGGATTCTGGAATCTGGCATCTGGCTTTTATAATGATGGGAAAAATGCAGAGCAGGTG 1677

QY	1694	TCCTTTCTCACCACCTGCCAATTGAGTCGGTGGATAAAGTCCCCAGCAACTGCTATGSCAAT	1755
Db	1678	TCCTTTAAATACCAATTGTTATAGAGTCTGTGGTGAATGTCCCGAAATTGCCATGGAAT	1737
QY	1754	GGTGACTGCATCTCTGGGACCTGCCACTGCTTCCCTGGGTTTCCCTGGGCCCCGACTGTGGC	1813
Db	1738	GGAGAAATGCGTTTCTGGAACTTGCCATTGTTTCCAGGATTTCTGGGTCCGGATTGTCTCA	1797
QY	1814	AGAGCCTCTGCCCCGCTCTGTAGCGGAAATGGCCAAATACATGAAAGGCAGATGCTTG	1873
Db	1798	AGAGCCGCTGTCCAGTGTATGTAGTGGCAACGGGCAAGTACTCCAAGGCGCGTGCCTG	1857
QY	1874	TGCCACAGTGGCTGGAAAGGCGCTGAGTGCAGTGTGCCACCAACCAAGTGTATCGATGTG	1933
Db	1858	TGTTTCAGCGGCTGGAAGGGCACCGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCA	1917
QY	1934	GCCTGCAGCAACCATTGGCACTCTGATCAGGGCACTTGCACTCTGCAACCTTGGCTACAAG	1993
Db	1918	CAGTGTGGGGTCTGTGGGATTTGTATCATGGCTCCCTGTGCTTGCAGCTCAGGATACAAA	1977
QY	1994	GGCAGAGACTGTGAGGAAGTGAAGTGCATGGACCCCACTGTTTTCAGGCCGGGTGTCTGC	2053
Db	1978	GGAGAAAGTTGTGAAGAAGCTGACTGTATAGACCCCTGGGTGTTCTAATCATGGTGTGTGT	2037
QY	2054	GTGAGAGGCGAATGCCATTGCTTTGTGGGATGGGAGGCACCAACTGCGAGACCCGAGG	2113
Db	2038	ATCCACGGGGAATGTCACTGCAGTCCAGGATGGGAGGTAGCAATTTGTAAATACTGAAG	2097
QY	2114	GCCACATGCTTAGACCAGTGTTCAGGCCACGGAACTTCTCTCCCGGACACCCGGGCTTTGC	2173
Db	2098	ACCATGTGTCCAGACCAGTGTCTCGGCCACGGAACTGTTCTCAAGAAAGTGGCTCTCTGC	2157
QY	2174	AGCTGTGACCCAAAGCTGACTGGACACGACTGTTCTATCGAGATCTGTGCTGCCGACTGT	2233
Db	2158	ACGTGTGACCCCTAACTGGACTGGCCAGACTGCTCAACGAAATATGTTCTGTGGACTGT	2217
QY	2234	GGTGGCCATGGCGTGTGCTAGGGGGCACCTGCCGTGCGAGGATGGCTGGATGGGGGCA	2293
Db	2218	GGCTCACACGGCGTTTGCAATGGGGGGACGTGTGCTGTGAAGAAGGCTGGACGGGCCCA	2277
QY	2294	GCCTGCGACACAGCGGGCCTGCAACCCCGCGCTGTCCGAGCATGGACCTGCCGCGACGGC	2353
Db	2278	GCCTGTAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCAGCGGACCTGCAAGGATGGC	2337
QY	2354	AAGTCGAGTGCAGCCCTGGTGGAAATGGCGAACACTGACCACTCGCTCACTATCTGGAT	2413
Db	2338	AAGTGTGAATGCAGCCAGGCTGGAATGGAGAGCACTGCACACTATCGCTCACTATTGGAT	2397
QY	2414	AGGGTAGT-----TAAAGAGGTTGCCCTGGTGTGCAATGGCAAC	2455
Db	2398	AAGATAGTTAAAGACAAGATAGGATATAAAGAGGGTGTCTGGTCTGTGCAACAGCAAT	2457
QY	2456	GGCAGATGTACCTTAGACCTGAATGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGA	2515
Db	2458	GGAAGATGTACCCCTGGACCAAAATGGCGGACATTTGTGTGTGCCAGCCCTGGATGGAGAGGA	2517
QY	2516	GCTGGCTGTGACACTTCCATGGAGACTGCCCTGGGTGACAGCAAGACAATGATGGAGAT	2575
Db	2518	GCAGGCTGTGACGTAGCCATGGAGACTCTTTTGCAACAGATAGCAAGGACAATGAAGGGGAT	2577
QY	2576	GGCCTGGTGGACTGCATGGACCTTGACTGTGCTGCCCTCAGCCCCCTGTGCCATATCAACCCG	2635
Db	2578	GGACTCATTTGACTGCATGGATCCCGATTGCTGCCCTACAGAGTTTCTGTGCCAGAAATCAGCCC	2637
QY	2636	CTGTGCCCTTGGCTCCCCTAACCTCTTGGACATCATCCAGGAGACACAGGTCTCTGTGTCTCA	2695
Db	2638	TATTGTCGGGACTGCCGGATCTCTCAGGACATCATTAGCCAAAAGCCTTCAATCGCCTTCT	2697
QY	2696	CAGCAGAACTACACTCTTCTATGACCGCATCAAGTTCTCTGTGGGACAGGACAGCACG	2755
Db	2698	CAGCAAGCTGCCAAATCCTTTTATGATCGAATCAGTTCAGTTTCTTATAGGATCTGTAGCACCC	2757

QY	2756	CACATAATCCCGGGAGAAACCCCTTTTGATGGAGGCATGCTTGTGTTATTCTGTGGCCAA	2811
Db	2758	CATGTTATACCTGGAGAAAGTCCTTTCAATAAGAGCCTTGTCATCTGTCTCATCAGAGGCCAA	2817
QY	2816	GTGATGACATCAGATGGAAACCCCTTGTTGGTGTGAACATCAGTTTTGTCAATAAACCTT	2875
Db	2818	GTACTGACTGCTGATGGAACTCCACTTATTGGAGTAAATGTCTCGTTTTTCCATTACCCA	2877
QY	2876	CTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTTGACTTGGTGACAAATGGCGGC	2935
Db	2878	GAATATGGATATACTATTAACCCGCCAGGACGGAATGTTTGACTTGGTGGCAATGGTGGG	2937
QY	2936	ATCTCCATCATCCTCGGTTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACCTGTGG	2995
Db	2938	GCCTCTCTAACTTTCGTATTGAAAGATCCCCATTCCTCACTCAGTATCATACTGTGTGG	2997
QY	2996	CTGCCATGGGATCGCTTCTTTTGTCTATGGAAACCATCATCATGAGACATGAGGAGAATGAG	3055
Db	2998	ATTCCATGGAATGTCTTTTATGTGATGGATACCTTAGTCATGGAGAAAGAGAGAATGAC	3057
QY	3056	ATTCCCAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTCGTCTCTCCATCCCCA	3115
Db	3058	ATTCCCAGCTGTGATCTGAGTGGATTCTGTGAGGCCAAATCCCATCATTTGTGTCTACCT	3117
QY	3116	CTGACGTCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCT	3175
Db	3118	TTATCCACCTTTTTCAGATCTTCTCCTGAAGACAGTCCCATCATTTCCCGAAACACAGGTA	3177
QY	3176	TTGCAGGAGGAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGG	3235
Db	3178	CTCCACGAGGAAACTACAATCCAGGAACAGATTTGAAACTCTCCTACTTGAGTCCAGA	3237
QY	3236	ACCCCTGGCTACAAATCTGTCTGAGGATCAGCTCACCCACCCGACCAATCCCTTCAAC	3295
Db	3238	GCTCAGGGTATAAGTCAGTTCTCAAGATCACCATGACCCAGTCTATTATTCATTTAAT	3297
QY	3296	CTCATGAAGGTGCACCTCATGGTAGCGGTGGAGGGCCGCTCTTTCAGGAAGTGGTTTCGT	3355
Db	3298	TTAATGAAGGTTTCATCTTATGGTAGCTGTAGPAGGAAGACTCTTCCAAAAGTGGTTTCCT	3357
QY	3356	GCAGCCCCAGACCTGTCTATTATTTCATTTGGGACAAGACAGACAGCTCTACAACCCAGAAG	3415
Db	3358	GCCTCACCAAACCTTGGCCTATACCTTTTCATATGGGATAAAAACAGATGCATATAATCAGAAA	3417
QY	3416	GTGTTTGGGCTTTCAGAAAGCCTTTGTTTCCGTGGGTTATGAATATGAATCCTGCCAGAT	3475
Db	3418	GTCTATGGTCTATCTGAAGCTGTTGTGTCAAGTGGATATGAGTATGAGTCGTGTTGGAC	3477
QY	3476	CTAATCCTGTGGGAAAAAGAACAAACAGTCTGCAGGGCTATGAAATTGACGGTCCAAG	3535
Db	3478	CTGACTCTGTGGAAAAAGAGGACTGCCAATTCGCAGGGCTATGAATTGGATCGTCCAAC	3537
QY	3536	CTTGGAGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTATCCTGCAC	3595
Db	3538	ATGGGTGGCTGGACATTAGATAAACATCACGTCTGGATGTACAGAACGGTATACTGTAC	3597
QY	3596	AAAGGGAATGGGGAGAACGAGTTTGTGTCTCAGCAGCCCTCCTGTCTATTGGGAGCATCATG	3655
Db	3598	AAGGGAACGGGAAAAACGAGTTTCATCTCCAGCAGCCTCCAGTCGTGAGTAGCATCATG	3657
QY	3656	GGCAATGGGCGCGGAGAAGCATCTCCTGCCCCAGCTGCAACGGCCCTTCTGTGAGCGCAAC	3715
Db	3658	GGCAATGGGCGAAGCGCAGCATTTCTCTGCCCCAGTTGCAATGGTCAAGCTGATGGTAAC	3717
QY	3716	AAGTCTCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCCTCTATGTGGGTGAT	3775
Db	3718	AAGTTACTGGCCCCAGTGGCGCTAGCTTGTGGGATCGATGGCAGTCTGTACGTAGGCGAT	3777
QY	3776	TTCAACTACATTAGAAAGGATCTTCCCCTCTGGAAATGTCAACCAACATCTCTAGAGCTGAGG	3835
Db	3778	TTCAACTACGTGCGGCGGATATTCCCTTCTGGAAATGTAAACAAGTGTCTTAGAACTAAGA	3837
QY	3836	AATAAAGATTTCAGACATAGTCACAGTCCAGCACACAAAATCTACTCTGGCCACAGACCCC	3895

Db 3838 AATAAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGGCAACGGATCCA 3897  
QY 3896 ATGAGTGGGGCCGCTCTTCCTTTCTGACAGCAACAGCCGGCGGTCTTTTAAATCAAGTCC 3955  
Db 3898 GTACGGGAGATCTGTACGTTCTGACACAAACACCCGACAGAAATTTATCGCCCAAAGTCA 3957  
QY 3956 ACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACCAG 4015  
Db 3958 CTTACGGGGGCAAAAGACTTGACTAAAAATGCAGAACTCGTCGACGGGACAGGGAGCAA 4017  
QY 4016 TGCCTCCCTTTTGATGACACTCGCTGCGGGATGGTGGGAAGGCCACAGAAAGCCACACTC 4075  
Db 4018 TGCCTTCCGTTTGACGAGCGAGATGTGGGATGGAGGAAGGCCGTGGAAAGCCACACTC 4077  
QY 4076 ACCAATCCCAGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACC 4135  
Db 4078 ATGAGTCCCAAAGGAATGGCAGTTGATAAGAAATGAATCAATCAACTCTTCTGGGCTTAACGATTTG 4137  
QY 4136 ATGATCAGACGCATCGATCAGAAATGGGATCAITCTCACCCCTGCTCGGCTCTAATGATCTC 4195  
Db 4138 ATGATTAGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTTAACGATTTG 4197  
QY 4196 ACATCAGCCCGCCACTCAGCTGTGATTCTGTCTATGGATATTTCAGGTAAGACTGGAG 4255  
Db 4198 ACTTCAGCCAGACCTTTAACTTGTGACACCAGCATGCACATCAGCCAGGTACGTCGGAA 4257  
QY 4256 TGGCCACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCTCGACAACAAT 4315  
Db 4258 TGGCCCACTGACCTAGCCATTAAACCTATGGATAACTCCATTATGTCTCGGATAATAAT 4317  
QY 4316 GTGGTCTCTGCAATCTCTGAAAAACCAACAGGTGGGCATTGTGCGCGGGAGGCCCATGCAC 4375  
Db 4318 GTAGTTTACAGATCACTGAAAATCGTCAAGTTCGATGCTGCTGGACGGCCCATGCAC 4377  
QY 4376 TGCCAGGTCCCTGGCATTTGACCACTTCCCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTG 4435  
Db 4378 TGTGAGTTCCCGGAGTGG---AATATCCTGTGGGGAAGCACGCGGTGCAGACAACTG 4434  
QY 4436 GAGTCAGCCACCGCTTTGGCTGTTTTCACACAATGGGGTCTCTGTATATTGCTGAGACTGAT 4495  
Db 4435 GAATCAGCCACTGCCATTGCTGTGCTCCTACAGTGGGTCTCTGTACATTACTGAAACTGAT 4494  
QY 4496 GAGAAAAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCT 4555  
Db 4495 GAGAAGAAAATTAACCGGATAAGGCAGGTCAACACAGATGGAGAAATCTCCTTAGTGGCC 4554  
QY 4556 GGGGCCCCAGTGGTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTCTGGAGAC 4615  
Db 4555 GGAATACCTTCAGAGTGTGACTGCAAAAAATGATGCCAACTGTGACTGTATACCAGAGTGA 4614  
QY 4616 GATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCCTTGGCTGTGTGCTGAT 4675  
Db 4615 GATGGCTACGCCAAGGATGCCAAACTCAGTGCCCATCTCCTCCCTGGCTGCTTCTCCAGAT 4674  
QY 4676 GGGGAGCTCTACGTGGCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAG 4735  
Db 4675 GGTACACTGTATATTGCAGATCTAGGGAATATCCGGATCCGGGCTGTGTCAAGAATAAG 4734  
QY 4736 CCTTTCCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCAGGAGCTCTAT 4795  
Db 4735 CCTTACTTAACTCTATGAACCTTCTATGAAGTTGGCTCTCCAACCTGATCAAGAACTCTAC 4794  
QY 4796 CTGTTTGTATACCCAGCAAGCACCTGTACACCCAAAGCCTGCCCAACAGGAGACTACCTG 4855  
Db 4795 ATCTTTGACATCAATGGTACTCAACCAATATACTGTAAGTTAGTCACTGGTATTACCTT 4854  
QY 4856 TACAACCTTCACTAGTGGGACGGCGACATCACTCACTCATCAGACAAACAATGGCAAC 4915  
Db 4855 TACAATTTTAGCTACGCAATGACAATGATATTAATTAATGCTGTGACAGACAGCAATGGCAAC 4914  
QY 4916 ATGGTAAATGTCCGCGGAGACTCTACTGGGATGCCCCCTCTGGCTGGTGGTCCCAGATGGC 4975

Db 4915 ACCCTTAGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCTCTGATAAC 4974  
QY 4976 CAGGTGTAAGTGGTACCACCATGGSCACCAACAGTGCACCTCAAGAGTGTGACCACACAAGGA 5035  
Db 4975 CAAGTGATATGTTTGACAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGA 5034  
QY 5036 CACGAGTTGGCCATGATGACATACCATGGCAATTCGCGCTTCTGGCAACCAAAAGCAAT 5095  
Db 5035 CTGGAAATAGTTTGTCTTACTTACCATGGCAATAGTGGCTTTTAGCCACTAAAAGTAT 5094  
QY 5096 GAAAACGGATGGACAAACATTTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACC 5155  
Db 5095 GAAACTGGATGGACAAACGTTTTTGGACTATGACAGTGAAGGTGCTGCTGACAAATGTTACG 5154  
QY 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAAGTATACAGACAGTTTCACTGATGTCAG 5215  
Db 5155 TTTCCAACTGGAGTGGTCAAAAACCTGCATGGGACATGGAACAAGGCTATCAAGTGGAC 5214  
QY 5216 GTAGAGACCTCCAGC---AAGGATGATGTCAACATAACCAACCAACCTGTCTGCCCTCAGGC 5272  
Db 5215 ATTGAGTCATCTAGCCGAGAGAAGATGTGAGCATCACTTCAAAATCTGTCTCTGATCGAT 5274  
QY 5273 GCCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGC 5332  
Db 5275 TCTTCTACACCATGGTTCAAGATCAGTTAAGAAAACAGTACCAAGATTGGTTATGACGGC 5334  
QY 5333 TCCTTGGCGTGTGTGGCAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTG 5392  
Db 5335 TCCCTCAGAAATTATCTACGGCAGTGGCCTGGACTCAGACTACCAACAGAGCCGACGTT 5394  
QY 5393 CTGGTGGCACCGTCAACCCACACCGTGGCAAGAGGAATGTCACTGTCGCCATCGACAAC 5452  
Db 5395 CTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAGCAAGCCCAAGGAAAGTCAATGTC 5454  
QY 5453 GGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGAGCGCTCGGGGCCAGGTCACTGTC 5512  
Db 5455 GGTCAAAACTTGGTGAATGGAGATTCCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTC 5514  
QY 5513 TTTGGCGCCGCTGCGGGTGCACAAACCGAAATCTCCTATCTCTGGACTTTTGATCGGTA 5572  
Db 5515 TTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTTTCAGTTGACTTTGATCGAACA 5574  
QY 5573 ACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCTGTACGAC 5632  
Db 5575 ACAAAGACAGAAAAGATCTATGACGACCACCGTAAATTTCTACTGAGGATCGCCTACGAC 5634  
QY 5633 CAGCGGGGGCGCCCAAGCTCTGGTCAACCCAGCAGCAGGCTGAATGGTGTCAACGTGACA 5692  
Db 5635 ACGTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGGCCGCTCAATGTCAAC 5694  
QY 5693 TACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAAAGATGGAA 5752  
Db 5695 TATTCAATCCACAGGTCAAATGCCAGCATCCAGCGAGGCCACCACTAGCGAGAAAGTAGAT 5754  
QY 5753 TACACACAGCGGGCGCATCACTCCAGGATCTTCGCTGATGGGAAGACATGGAGCTAC 5812  
Db 5755 TATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTGTGATGGTAAACATGGAGTTAC 5814  
QY 5813 ACATATTAGAGAAAGTCCATGGTCTGCTACTACAGCCAGAGGCGAGTATATCTTTGAG 5872  
Db 5815 ACATATTAGAAAAGTCCATGGTCTCTCGGGTCTTGTGATGGTAAACATGGAGTTAC 5874  
QY 5873 TTGCAACAAGAAATGACCCGCTCTCTTGTGACGATGTCACCAACCTGTCGCGGCGAGACTA 5932  
Db 5875 TAGCATATGTGGACCCGCTGTCTGCCATCACCATGCCCAGTGTGGCTCGCCACACCATG 5934  
QY 5933 GAGACCATCCGCTCAGTGGGCTACTACAGAAAACATCTATCAGCCCCCTGAGGGCAATGCC 5992  
Db 5935 CAGACCATCCGATCCATTTGGCTACTACCGCAACATATACAACCCCCCGGAAAGCAACGCC 5994  
QY 5993 TCAGTCATACAGGACTTCACTGAGGATGGGCACCTCTTACACCTTCTACCTGGGCACT 6052  
Db 5995 TCCATCATCACGGACTACACGAGGAAGGGCTGCTTCTACAAACAGCTTTCTTGGGTACA 6054

QY	6053	GGCCG	CAGG	TGAT	AATAC	AAAGT	ATGG	CAAACT	GTCA	AAAGCT	TGGC	AGAG	ACGCT	CTAT	GAC	6111
DB	6055	AGTCG	AGGG	TCTAT	TCAA	ATAC	AGAA	GGCAG	ACTAG	GGCTCT	TCAG	AAATTT	TATAT	GAT	6114	
QY	6113	ACCAC	CAAG	GTCAG	TTTCA	CTAT	GAC	GAGAC	GGCAG	GCCAT	GTGA	AGAC	CAAC	CACTA	6172	
DB	6115	AGCACA	AGAGT	CAAGTTT	TACCT	TATG	ATGA	AAAC	AGCAG	GAGTCT	TAAAG	CACAG	TAAC	CACTC	6174	
QY	6173	CAGA	ATGAG	GGCTT	CAC	CTGC	ACC	CACTCC	GGCTAC	CGTAC	GATTTG	GGCC	CCCTG	ATTGAC	6232	
DB	6175	CAGAG	TGAT	GGTTTT	TATTG	CACCAT	TAG	ATAC	AGCA	AAATGGT	CCCC	CTGAT	TGAC	AGG	6234	
QY	6233	CAGAT	CTTCC	GGCTT	CAC	TGAG	GAAG	GCAT	GGTCA	ACG	CCCGTTT	TGAC	TAC	AACTAT	6292	
DB	6235	CAGAT	TTTCC	GGCTT	TAGT	GAAG	TGGG	ATGG	TAAATG	CAAG	ATTTG	ACTAT	ATAG	CATGAC	6294	
QY	6293	AACAG	CTTCC	GGGTG	ACC	AGCAT	GC	AGGCTGT	GTATCA	ACG	AGAC	CCCC	CACTG	CCCAT	6352	
DB	6295	AACAG	CTTTC	GAGT	GAC	AGCAT	GC	AGG	TGTGAT	CAATG	AAAC	CGCC	CACTG	CCCAT	6354	
QY	6353	CTCTAT	CGTAT	GTAT	GTGT	CA	GGCA	AGAC	AGAG	CAGTTT	TGGG	AAAGTTT	TGGT	GTCT	6412	
DB	6355	CTGTAT	CAGTTT	GAT	GACAT	TTCT	TG	CAAA	AGTTG	AGCAG	TTT	TG	AAAGTTT	TGG	6414	
QY	6413	TACTAT	GACAT	TAA	CCAG	ATCAT	CAC	CACAG	CTGT	CATG	ACCC	CACAC	CAAG	CACTTT	6472	
DB	6415	TATTAT	GATAT	TAA	CCAG	ATCAT	TTCT	TAC	AGCTGT	AATG	ACCT	TATAC	GAA	GCAT	6474	
QY	6473	GCATAT	GGC	AGATGA	AGAA	GTGC	AGTAT	GAG	ATCTT	CCGCT	CGCT	CA	TGTA	CTGG	6532	
DB	6475	GCTCAT	GGCC	GTATCA	AGG	AGAT	TTCA	ATAT	GAG	ATAT	CAGG	TCGCT	CAT	GTACT	6534	
QY	6533	ACCGT	CCAG	TATG	ATA	CAT	TGGG	CGAGT	AGT	AGAA	GGAG	CTGA	AGG	TAGG	6592	
DB	6535	ACAAT	TCAG	TATG	ATA	CAT	GGT	CGGGT	AA	CCA	AGAG	AGAT	TAA	ATAGG	6594	
QY	6593	GCCAA	TACCA	CTCG	TACT	CTCT	ATG	AGTAT	GTAT	GCTG	ACG	GGCC	AGCTG	CAGAC	6652	
DB	6595	GCCAA	CACCA	CCAAT	TGCTT	ATG	AAAT	ATGAT	GTGAT	GGAC	AGCT	CCAA	ACAG	TTTAC	6654	
QY	6653	ATCAAT	GACA	AGCC	ACTCT	CTG	CGCT	TAC	AGCTAC	AGCT	CAAT	GGG	AACTG	CACTT	6712	
DB	6655	CTCAAT	GAAA	AGATA	TCT	TG	CGGT	TACA	ACTAC	GATCT	GAA	TG	AAACCT	CCAT	6714	
QY	6713	AGCCCT	GGAA	CAGT	GC	AGGCT	CAC	ACC	ACTAC	GGTAT	GACAT	CCG	CAC	CGCAT	6772	
DB	6715	AACCC	AAAGTAA	CAGT	GGCGT	CTG	AC	CCCCCT	T	CGTAT	GACCT	G	GAGAC	AGAT	6774	
QY	6773	CGGCT	GGGTG	ACGTG	CA	ATACA	AGAT	TGGAT	TGGAT	TGGCTT	CTCT	GAGG	CAGCG	GGCGGT	6832	
DB	6775	CGACT	GGGTG	ATGTT	CA	ATAT	CGTT	TGGAT	GA	AGAT	GTGTT	CTC	ACGT	CAA	6834	
QY	6833	GATAT	CTTT	TGAGT	ACA	ACTC	AGCT	TGGCCT	GTCT	CATCA	AGCC	TACA	CCGGG	CTGG	6892	
DB	6835	GAAAT	CTTT	TGAAT	ATAG	CTCCA	AGGG	CTTCT	AACT	CGAG	TTTAC	AGTAA	AGG	CACTG	6894	
QY	6893	TGGAG	TGTC	AGGTAC	CCGTAC	GAT	TGGCCT	TGGG	CGCGGT	GTCT	CCAG	CAAG	AGAG	CAGCC	6952	
DB	6895	TGGAC	AGT	GATCT	AC	CGGTAT	GAC	GGCCT	TGGG	AAAGG	CGGTGTTT	CTAG	CA	AAAC	6954	
QY	6953	AGCC	ACC	ACCTG	CAGT	TTCTT	ATG	CAGAC	CTTG	ACCA	CCCC	CAAGG	TCA	CCCC	7012	
DB	6955	GGAC	AGCAC	CTG	CAGTTT	TTTT	TAT	GTGCT	GACTTAA	CTTAT	TCC	CACTAG	GATTAC	TATGTC	7014	
QY	7013	TACA	ACC	ACTCC	AGCT	CTG	AGAT	CACCT	CCCTCT	ACTAC	GA	CTTG	CAAG	GACAC	7072	
DB	7015	TACA	ACC	ATTC	GAGT	TC	AGAA	TTAC	CTCC	CTGT	AT	TATG				

QY	7133	CTTGCTGCTTTTAGTGAACAGGTTTGTATGATCAAGCAAAATCCTGTACACAGCCTATGGG	7199
DB	7135	CTGGCTGTGTTAGTAGCAATGGGCTTATGTGTAAGACAGATTTCAGTACACTGCATATGGG	7194
QY	7193	GAGATCTACATGATACCAACCCCAACTTTTTCAGATCATCATAGGCTACCATGGTGGCCTC	7252
DB	7195	GAAATCTATTTTGACTCTAATATTGACTTTCAACTGGTAATTGGATTTCATGGTGGCCTG	7254
QY	7253	TATGATCCACTCACCAAGCTTGTCCACATGGSCCGCGAGATTATGATGTGCTGGCCGGA	7312
DB	7255	TATGACCCACTCACCAAAATTAAATCCACTTTGGAGAAAGAGATTATGACATTTTGGCAGGA	7314
QY	7313	CGCTGGACTAGCCAGACCAAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCT	7372
DB	7315	CGGTGGACAACACTGACATAGAAATCTGGA--AAAGAAATGGGAAGGACCCAGCTCCT	7371
QY	7373	TTTAATCTCTATATGTTCAAAAACAACCCCATCAGCAACTCCCAGGACATCAAGTGC	7432
DB	7372	TTTAACTTGTACATGTTTAGGAATAACAACCCCTGCAAGCAAAATCCATGACGTGAAAGAT	7431
QY	7433	TTCATGACAGATGTTAAACAGCTGGCTGCTCACTTTGGATTCCAGCTACACAACCTGATC	7492
DB	7432	TACATCACAGATGTTAAACAGCTGGCTGGTGACATTTGGTTTCCATCTGCACAATGCTATT	7491
QY	7493	CCTGGTTATCCCAAACCAAGACATGGATGCCATGGAAACCTCCTACGAGCTCATCCACACA	7552
DB	7492	CCTGGATTCCCTGTGCCAAATTTGATTTTAAACAGAACTTCTTTACGA-----A	7539
QY	7553	CAGATGAAAACGGCAGGAGTGGGACAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTA	7612
DB	7540	CTTGTGAAGAGTCAGCAGTGGGATGATATACCGCCCATCTTCGGAGTCCAGCAGCAAGTG	7599
QY	7613	CAGAAAGCAGTCAAGGCCCTTTGTACCTTAGAACCGTTTGTACCAGCTCTATGGCTCCACA	7672
DB	7600	GCGGGCAGGCCAAGGCCTTCTCTGCTGCTGGGGAAGATGCCGAGGTGCAGTGCAGCCGG	7659
QY	7673	ATCACACAGCTGCCAGCAGGCTCCAAAGACCAGAAAGTTTGCATCCAGCGGCTCAGTCTTT	7732
DB	7660	CGCCGGGCCGGC--GGCGCGCAGTCCCTGGCTGTGGTTGCGCACGGTCAAGTCCGCTGATC	7716
QY	7733	GGCAAGGGGTCAGTTTGCTTGCCTTGAAGGATGGCCGAGTGACCACAGACATCATCAGTGTG	7792
DB	7717	GGCAAGGGCGTCAATGCTGGCCGTGAGCCAGGGCCGCGTGCAGACCAACCTGCTCAACATC	7776
QY	7793	GCCAAATGAGGATGGCCAAAGGTTTGCTGCCATCTTGAACCATGCCCACTACCTAGAGAAC	7852
DB	7777	GCCAAAGGAGTCAATCAAGTGGCGCGCGTGTCTCAACAACGCCTTCTACCTGGAGAAC	7836
QY	7853	CTGCACTTCACCAATGATGGGGTGGATACCCATTACTTTGTGAAACCAGGACCTTCAGAA	7912
DB	7837	CTGCACTTCACCATCGAGGCAAGGACACCGCACCTTCTCATCAAGACCAACCGCCCGAG	7896
QY	7913	GGTGACCTGGCCATCCTGGGCCTCAGTGGGGGGCGCGAACCCCTGGAGAATGGGGTCAAC	7972
DB	7897	AGCGACCTGGGCACGCTGCGGTTGACCAGCGGCCGCAAGGCGCTGGAGAACCGCATCAAC	7956
QY	7973	GTCACCTGTGCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATC	8032
DB	7957	GTGACGGTGTGCGAGTCCACCACGGTGGTGAACCGGCAAGGCGCTGGAGAACCGCATCAAC	8016
QY	8033	CAGCTCCAGTACGGGGCACTGTGCTTGAACACACGCTACGGGACAACGTTTGGATGAGGAG	8092
DB	8017	GAGATGCAGTTTCGGCGCGCTGGCGCTGCACGTTGCGCTACGGCATGACCCCTGGACGAGGAG	8076
QY	8093	AAGGCACGGGTCTGTAGACTGGCCCGCAGAGAGCCGTGCGCCCAAGCGTGGGCCCGCGAG	8152
DB	8077	AAGGCGCGCATCCTTGGAGCAGGCGCGGACGGCGCGCTGCCCCGGGCTTGGGCGCGCGAG	8136
QY	8153	CAGCAGAGACTGCGGGAAGGGGAGGAAGGCCCTGCGGGCCTTGGACAGAGGGGGAGAACGAG	8212
DB	8137	CAGCAGCGCGTGCAGCAGCGCGAGGAGGGCGCGCTCTGGAACGAGGGCGGAGAACCGG	8196
QY	8213	CAGGTGCTGAGCACAGGCGCGGTGCAAGGCTACAGCGCTTTTTCGTGATCTCTGTCGAG	8272

Db 8197 CAGCTGCTGAGCGCGCAAGTGCAGGGCTACGACGGGTACTACGTACTCTCGGTGGAG 8256  
QY 8273 CAGTACCAGAACTGTACAGACGGCCCAACAACATCCACTTCATGAGACAGAGCGAGATG 8332  
Db 8257 CAGTACCCCGAGCTGGCCGACAGCGCCCAACAACATCCAGTTCCTGCGGCAGAGCGAGATC 8316  
QY 8333 GGCAGGAGGTGAC 8345  
Db 8317 GGCAGGAGGTAAC 8329

RESULT 7  
US-10-038-854-39  
; Sequence 39, Application US/10038854  
; Publication No. US20040022781A1

; GENERAL INFORMATION:  
; APPLICANT: Spytex, Kimberly A  
; APPLICANT: Li, Li  
; APPLICANT: Wolenc, Adam R  
; APPLICANT: Vernet, Corine  
; APPLICANT: Eisen, Andrew J  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Gorman, Linda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Guo, Xiaojia S  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Erik  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: MacDougall, John R  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-230  
; CURRENT APPLICATION NUMBER: US/10/038,854  
; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: 60/258,928  
; PRIOR FILING DATE: 2000-12-29  
; PRIOR APPLICATION NUMBER: 60/259,415  
; PRIOR FILING DATE: 2001-01-02  
; PRIOR APPLICATION NUMBER: 60/259,785  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 60/269,814  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/279,832  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,833  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,863  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/283,889  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/284,447  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/286,683  
; PRIOR FILING DATE: 2001-04-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 411  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 8473

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-038-854-39  
Query Match 41.0%; Score 3422; DB 17; Length 8473;  
Best Local Similarity 67.1%; Pred. No. 0;  
Matches 5027; Conservative 0; Mismatches 2370; Indels 93; Gaps 9;  
QY 880 CGGGCACTTCCTCTTCAAGCCTGG--AGGCACCTCCCGCTCTTCTGCACCACATCACC 936  
Db 725 CAGGCATTTCTTATTCAAAACAGGAACAGGTACAACGCCACTGTTTCAGTACTGCAACCCC 784  
QY 937 AGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCAG 996  
Db 785 AGGATACACAATGGCATCTGGCTCTGTTTATTACCACTACTCGGCCACTACCTAGAAA 844  
QY 997 CACCTTCGCCCGCGCGCTTTAAACCTCAAGAAAGCCCTCAAGTACTGTAACTGGAAGTG 1056  
Db 845 CACCCATCAAGAAGTGTCTTTAAATTCAAGAAGTCTTCAAGTACTGTAGCTGGAATG 904  
QY 1057 CGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGTGTCATCTCTGTCATCTTGT 1116  
Db 905 CACTGCACTGTGTCCGTAGGGTCTCGGTCTCTCTGGAATACTCTCTGCTTATTAT 964  
QY 1117 GGCCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGATGA 1176  
Db 965 AGCATGCATCTCTTTGGCCCTCAACTGGCAGCTACAGCAGACTGAAAATGACACATTGA 1024  
QY 1177 GATCAGGAGGACACAGCCAGCTTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTC 1236  
Db 1025 GA-----ATGGAAAAGTGAATCTGTATACCATGCCAACAAACACTGTGTCTATTCCCTTC 1078  
QY 1237 AGGGGCACCTGGCTTAGAGACCCCTGACAGGAAAGGAAAGGAAACACAGAAAGAAAGCC 1296  
Db 1079 TGGAGACA-----ATGGAAAATT 1096  
QY 1297 CAGTAGTTTCTTTCCAGAGGACAGTTTCATAGATTCTGGAGAAATTCATGTTGGGAGGCG 1356  
Db 1097 AGGTGATTTACGCAAGAAAATAACCATAGATTCCGGAGAACTTGATATTGGCCGAAG 1156  
QY 1357 AGCTTCCCAGAAAGATTCTCTCGGCACCTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416  
Db 1157 AGCAATTCAAGAGATTCTCTCGGGATCTTCTGGAGATCACAGCTCTTTCATTGATCAGCC 1216  
QY 1417 TGTGCATCTGAAATTCATGTGTCTCTTGGGAAAGGACGCTCTGTCATTTATGGCAG 1476  
Db 1217 ACAGTTTCTTAAATTCAATATCTCTTTCAGAAAGGATGATGATGATGATGATGATGATG 1276  
QY 1477 AAAAGCCTCCCTCTCTACATACACAGTTTGAATTTGTGGAGCTGCTGGATGGCAGGAG 1536  
Db 1277 GAAAGCCTTACCGCTTCCCATACTCAGTATGACTTCGTGGAGCTCCTGGATGGCAGCAG 1336  
QY 1537 GCTCCTAACCCAGGAGCGCGGAGCCTTAGAGGGGACCCCGCCAGTCTCGGGGAACTGT 1596  
Db 1337 GCTGATTGCCAGAGAGCAGCGGAGCCTGCTTGAGACGAGAGAGCGCGGCGGCGGCGAG 1396  
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTT 1656  
Db 1397 ATCGTCAAGCTTCATGAGCGCGCTTTATCCAGTACTTGGATTCTGGAATCTGGCATCT 1456  
QY 1657 GGCTTTTACAATGACGGAAGGAGTCAAGAGTGGTTTCTTCTCACCCTGCCATTGA 1716  
Db 1457 GGCTTTTATAATGATGGGAAAATGCAGAGCAGGTTCTTTTAAATACCATTTGTTATAGA 1516  
QY 1717 GTCGTGGATAACTGCCCGCAACTGCTATGGCAATGGTGAATGCTCTCTGGGACCTG 1776  
Db 1517 GTCGTGGTGAATGTCCCGAAAATTGCCAATGGAAATGGAGAAATGCGTTTCTGGAACTTG 1576  
QY 1777 CCACTGCTTCTGGGTTTCTGGGCCCCGACACTGTGGCAGAGCCTCTCTGCCCGCTGCTG 1836  
Db 1577 CCATTGTTTCCAGGATTTCTGGGTCCGGATTGTTCAAGAGCGCGCTGTCCAGTGTATG 1636  
QY 1837 TAGCGGAAATGGCCAATACATGAAAGGACAGATGCTTGTGCCACAGTGGCTGGAAGGCGC 1896

Db 1637 TAGTGGCAACGGGAGTACTCCAAAGGCGCTGCCTGTGTTTCAGCGGTGGAAGGGCAC 1696  
 QY 1897 TGAGTGGGATGTGCCCCACCAACCAAGTGTATCGATGTGGCCCTGCAGCAACCATGGCACCTG 1956  
 Db 1697 CGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGGTCTGTGGGATTG 1756  
 QY 1957 CATCACGGGCACTGCATCTGCAACCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGA 2016  
 Db 1757 TATCATGGGCTCTTGTGCTTGCACCTCAGGATACAAAGGAAAGTTGTGAAGAGCTGA 1816  
 QY 2017 CTGCATGGACCCACATGTTTCAGGCCGGGGTGTCTGCGTGAGAGCGGAATGCCATTGCTT 2076  
 Db 1817 CTGTATAGACCCCTGGGTGTTCTAATCATGGTGTGTATCCACGGGGAATGTCACTGCAG 1876  
 QY 2077 TGTGGGATGGGAGGCACCAACTGCGAGACCCCGAGGCCACATGCTTAGACCAGTGTTC 2136  
 Db 1877 TCCAGGATGGGAGGTAGCAATTGTGAAATACTGAAGACCATGTGTCCAGACCAGTGTCTC 1936  
 QY 2137 AGCCACCGAACCTTCTCCCGGACACCGGGCTTTGACGCTGTGACCCCAAGCTGGACTGG 2196  
 Db 1937 CGGCCACGGAACGTATCTTCAAGAAAGTGGCTCCTGCACGTGTGACCCCTAAGTGGACTGG 1996  
 QY 2197 ACACGACTGTTCTATCGAGATCTGTGTCGCCGACTGTGGTGGCCATGGCGTGTGCGTAGG 2256  
 Db 1997 CCCAGACTGCTCAAACGAAATATGTTCTGTGACTGTGGCTCAGCGCGTGTTCATGGG 2056  
 QY 2257 GGGCACCTGCCGTGCGAGGATGGCTGGATGGGGGAGCCCTGCGACCGAGCGGGCTGCCA 2316  
 Db 2057 GGGGACGTGTGCTGTGAAGAGGCTGAGCGGGCCAGCCCTGTAATCAGAGAGCCCTGCCA 2116  
 QY 2317 CCCGCGCTGTCCGAGCATGGGACCTGCGCGGACGGCAAGTGGAGTGCAGCCCTGGCTG 2376  
 Db 2117 CCCCCGCTGTCCGAGCACGGGACCTGCAAGGATGGCAAGTGTGAATGCAGCCAGGGCTG 2176  
 QY 2377 GAATGGCGAAGTGCACCATCGCTCACTATCTGGATAGGGTAGT----- 2421  
 Db 2177 GAATGGAGAGCACTGCACTATCGCTCACTATTTGGATAAGATAGTTAAAGACAAGATAGG 2236  
 QY 2422 ---TAAAGAGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGATGTACCTTAGACCTGAA 2478  
 Db 2237 ATATAAGAGGGTGTCTCTGGTCTGTGCAACAGCAATGGAAGATGTACCTTGAGACCAAAA 2296  
 QY 2479 TGGTTGGCACTGGCTCTGCCAGCTGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGA 2538  
 Db 2297 TGGCGGACATTGTGTGTGCCAGCCTGGATGGAGAGGAGCAGGCTGTGACGTAGCCATGGA 2356  
 QY 2539 GACTGCCCTGCGGTGACAGCAAGACAAATGATGGAGATGGCCTGGTGGACTGCATGGACCC 2598  
 Db 2357 GACTCTTTGCACAGATAGCAAGGACAAATGAAGGGGATGGACTCATTTGACTGCTATGATCC 2416  
 QY 2599 TGACTGTGCTCCAGCCCTGTGTCCTATCAACCCGCTGTGCTTGGCTCCCTTAACCC 2658  
 Db 2417 CGATTGCTGCCTACAGAGTTCTCTGCCAGATCAGCCCTATTGTGCGGGACTGCCGGAACC 2476  
 QY 2659 TCTGGACATCATCCAGGAGACACAGGTCCTGTGTCAAGCAGCAACCTACACTCCTTTCTA 2718  
 Db 2477 TCAGGACATCATTAGCCAAAGCCTTCAATGCGCTTCTCAGCAAGCTGCCAAATCCTTTTA 2536  
 QY 2719 TGACCGCATCAAGTTCTCGTGGGAGGACAGCAGCAGCACAATAATCCCGGGGAGAACCC 2778  
 Db 2537 TGATCGAATCAGTTTCTTATAGGATCTGTATAGCACCCTATGTTATACCTGGAGAAAGTCC 2596  
 QY 2779 CTTTGTAGGAGGGATGCTTGTGTTATTCTGTCGCAAGTGTATGACATCAGATGGAACCC 2838  
 Db 2597 TTTCAATAAGAGCCTTGCATCTGTCTCAGAGGCCAAGTACTGACTGCTGATGGAACCTCC 2656  
 QY 2839 CCTGGTTGGTGAACATCAGTTTGTGCAATAACCTCTCTTTTGGATATACAATCAGCAG 2898  
 Db 2657 ACTTATTGGAGTAATGTCTCGTTTCTCCATTACCAGAAATATGGATATACTATTACCCG 2716  
 QY 2899 GCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTCCCATCATCTCGGTTTCCA 2958

Db 2717 CCAGGACGGAATGTTTGACTTGTGGCAAAATGGTGGGGCCCTCTCTAACTTTGGTATTGA 2776  
 QY 2959 GCGGGCACCTTTTCATCACACAGAGACACACCCTGTGGCTGCCATGGGATCGCTTCTTGT 3018  
 Db 2777 ACGATCCCATTCCTCACTCAGTATCATACTGTGTGGATTCATGGAATGTCTTTTATGT 2836  
 QY 3019 CATGGAACCATCATCATGAGACATGAGGAGAAATGAGATTCCAGCTGTGACCTGAGCAA 3078  
 Db 2837 GATGGATACCTAGTTCATGGAGAAAGAGAAATGACATTCCAGCTGTGATCTGAGTGG 2896  
 QY 3079 TTTTGGCCGCCCAACCCAGTCTCTCTCCATCCCACTGACGTCCTTCGCCAGCTCCTG 3138  
 Db 2897 ATTCTGTAGGCCAAATCCCATCATTTGTGTCTATCACCTTTTATCCACCTTTTCAGATCTC 2956  
 QY 3139 TGCAGAGAAAGGCCCCCATTTGTGCCGGAATTCAGGCTTTTGAGGAGGAAATCTCTATCTC 3198  
 Db 2957 TCCTGAAGACAGTCCCATCATTTCCCGAAACACAGGTACTCCAGGAGGAAACTACAATTC 3016  
 QY 3199 TGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCCGACCCCTGGCTACAAATCTGTCT 3258  
 Db 3017 AGGAACAGATTGAAACTCTCTCTACTTGTGATTCAGAGCTGCAGGGTATAAGTCAGTTCT 3076  
 QY 3259 GAGGATCAGCCTCACCCACCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGCT 3318  
 Db 3077 CAAGATCACCATGACCCAGTCTATTATTCCATTAAATTTAATGAAGGTTCATCTTATGT 3136  
 QY 3319 AGCGGTGGAGGGCCGCTCTTTCAGGAAGTGTTCGCTGCAGCCCGACACCTGTCTATTTA 3378  
 Db 3137 AGCTGTAGTAGGAAGACTCTTCCAAAAGTGGTTTCTCTGCCCTCACCAACTTGGCCCTATAC 3196  
 QY 3379 TTTTCAATTTGGACAAGACAGACGCTCTACAAACAGAGGTGTGTGGGCTTTTCAGAAGCCTT 3438  
 Db 3197 TTTTCATATGGGATAAAACAGATGCATATAATCAGAAAAGTCTATGGTCTATCTGAAGCTGT 3256  
 QY 3439 TGTTCCTGGTGTATGAATATGAATCTCTGCCAGATCTAATCTGTGGGAAAAAAGAAC 3498  
 Db 3257 TGTGTCAAGTTGGATATGATATGATCGTCTGTTGGACCTGACTCTGTGGGAAAAAGAGAC 3316  
 QY 3499 AACAGTGTGCAGGGCTATGAAATTCACGCGTCCAAGCTTGGAGGATGGAGCCTTAGACAA 3558  
 Db 3317 TGCCATTCTGCAGGGCTATGAATTTGGATCGCTTCCACATGCGTGGCTGGACATTAGATAA 3376  
 QY 3559 ACATCATGCCCTCAACATTCAAAGTGGTATCTCTGCACAAAGGAAATGGGAGAACCCAGTT 3618  
 Db 3377 ACATCAGTGTGGATGTACAGAACGGTATCTGTACAGGAAACCGGGGAAAAACCCAGTT 3436  
 QY 3619 TGTGTCTCAGCAGCCTCCTGTCTATTGGAGCATCATGGGCAATGGGCGCCCGAGAGCAT 3678  
 Db 3437 CATCTCCAGCAGCCTCCAGTCTGTAGTAGCATCATGGGCAATGGGCGAAGGCGCAGCAT 3496  
 QY 3679 CTCCTGCCCGAGTGCACCGGCTTGTGACGCGCAACCAAGCTCTGGCCCCCAGTGGCCCT 3738  
 Db 3497 TTCTTGCCCCAGTTGCAATGGTCAAGCTGATGGTAAACAAGTTACTTGGCCCCCAGTGGCGCT 3556  
 QY 3739 CACCTGTGGCTCTGACGGGAGCCTCTATGTGGTGTATTTCAACTACATTAGAAGGATCTT 3798  
 Db 3557 AGCTTGTGGGATCGATGGCAGTCTGTACTAGGCGGATTTCAACTACGTCGGCGGATATT 3616  
 QY 3799 CCCCTCTGGAATGTCAACCAACATCTTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCA 3858  
 Db 3617 CCCTTCTGGAATGTAAACAAGTCTTTAGAACTAAGAAATAAAGATTTTAGACATAGCAG 3676  
 QY 3859 CAGTCCAGCACCAATACTACCTGGCCACAGACCCCATGAGTGGGGCCGCTCTCTCCTTTC 3918  
 Db 3677 CAACCCAGCTCATAGATACTACCTTTCACCGGATCCAGTCAAGGAGATCTGTACGTTTC 3736  
 QY 3919 TGACAGCAACAGCCGGGGTCTTTTAAATCAAGTCCCACTGTGTGGTGAAGGACCTTGT 3978  
 Db 3737 TGACACAAACACCCGAGAAATTTATCGCCCAAGTCACTTACGGGGGCAAGACTTGAC 3796  
 QY 3979 CAAGAACTCTGAGGTGGTTGGGGGACAGGTGACAGTGCCTCCCTTCCCTTTGATGACACTCG 4038  
 Db 3797 TAAAAATGCAGAGTCTGTCGAGGACAGGGGAGCAATGCCCTTCCGTTTACGAGGCGGAG 3856









Db 4051 CAGCCAGGTACGCTCGGAATGGCCCACTGACCTAGCCATTAAACCCTATGGATAACTCCAT 4110  
Qy 4297 TTATGTCTCGACAAAGTGTGGTCTCTGCAAAATCTCTGAAATCCACAGGTGCGCATTTGT 4356  
Db 4111 TTATGTCTCGATAAATAATGTAGTTTACAGATCACTGAAATCGTCAAGTTCGCATTGC 4170  
Qy 4357 CGCCGGGAGGCCCATGCACTGCCAGGTCCTGGCAATTGACCACCTTCCTGCTAAGCAAGGT 4416  
Db 4171 TGCTGGACGGCCCATGCACTGTGAGTTCCCGAGTGG--AATATCCTGTGGGGAAGCA 4227  
Qy 4417 GGCCATCCACGCAACCTCGGAGTCAAGCCACCGCTTTGGCTGTCTTTCACACAAATGGGTCCT 4476  
Db 4228 CGCGGTGCAGACAACTCTGGAATCAGCCACTGCAATGCTGTCTCTACAGTGGGTCCT 4287  
Qy 4477 GTATATTGCTGAGACTGATGAGAAAAGATCAACCGCATCAGGCAGGTCAACCACTAGTGG 4536  
Db 4288 GTACATTACTGAAACTGATGAGAGAAAATTAACCGGATAAGGCAGGTCAACACAGATGG 4347  
Qy 4537 AGAGATCTCACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTG 4596  
Db 4348 AGAAATCTCCTTAGTGGCCGGAATACCTTCAGAGTGTGACTGCAAAAATGATGCCAACTG 4407  
Qy 4597 TGATTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTC 4656  
Db 4408 TGACTGTTACAGAGTGGAGATGGCTACGCCAAGGATGCCAAACTCAGTCCCCCATCTC 4467  
Qy 4657 CTTGGCTGTGTGCTGATGGGGAGCTCTACGTGGCCGACCTTGGGAAACATCCGAATTGC 4716  
Db 4468 CCTGGCTGCTTCTCCAGATGGTACACTGTATATTGCAGATCTAGGGAATATCCGGATCCG 4527  
Qy 4717 GTTTATCCGGAAGAACAGCCTTTCCCTCAACACCCAGAACATGTATGAGCTGTCTTCACC 4776  
Db 4528 GGCTGTGTCAAAGAAATAGCCCTTTACTTAACTTATGAACCTTCTATGAAGTTGCGTCTCC 4587  
Qy 4777 AATTGACCAGGAGCTTATCTGTTTGTATACCAACCGGAAGCACCTGTACACCCAAAGCCT 4836  
Db 4588 AACTGATCAAGAACTCTACATCTTTGACATCAATGGTACTCACCAATATACTGTAAGTTT 4647  
Qy 4837 GCCCACAGGAGACTACCTGTACAACCTTCCTCAACACCCAGAACATGTATGAGCTGTCTTCACC 4896  
Db 4648 AGTCACCTGGTGATTACCTTTACAATTTTAGCTACAGCAATGACAATGATATTACTGCTGT 4707  
Qy 4897 CACAGACAACAATGGCAACATGGTAAATGTCCGCGGAGACTCTACTGGGATGCCCTCTG 4956  
Db 4708 GACAGACAGCAATGGCAACACCCCTTAGAATTAGACGGGACCCAAATCGCATGCCAGTTGC 4767  
Qy 4957 GCTGGTGGTCCCAGATGCGCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCACCTCAA 5016  
Db 4768 AGTGGTGTCTCCTGATAACCAAGTGATATGGTTGACATAGGAACAATGGATGTTTGAA 4827  
Qy 5017 GAGTGTGACCAACACAGGACACGAGTTGGCCCATGATGACATACCATGGCAATTCCGGCCT 5076  
Db 4828 AGGCATGACTGCTCAAGGACTGGAATTAGTTTTGTTTACTTACCATGGCAATAGTGGCCT 4887  
Qy 5077 TCTGGCAACCAAAAGCAATGAAAACGGATGGACAAACATTTTATGAGTACGACAGCTTTGG 5136  
Db 4888 TTTAGCCACTAAAAGTGATGAAACTGGATGGACAAACGTTTTTTTGACTATGACAGTGAAG 4947  
Qy 5137 CCGCCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTTCCGAAGTGATACAGA 5196  
Db 4948 TCGTCTGACAAATGTTACGTTTCCAACTGGAGTGGTCAAAAACCTGCATGGGGACATGGA 5007  
Qy 5197 CAGTTCAGTGCATGTCAGGTAGAGACCTCCAGC--AAGGATGATGTCAACCATCAACAC 5253  
Db 5008 CAAGGCTATCACAGTGGACATGAGTCACTAGCCGAGAAGAAGATGTACGATCACTTC 5067  
Qy 5254 CAACCTGTCTGCCCTCAGGCGCCTTCTACACACTGTGTGCAAGACCAAGTCCGGAACAGCTA 5313  
Db 5068 AAATCTGTCTCGATCGAFTCTTTCTACACCATGGTTCAAGATCAGTTAAGAAAACAGCTA 5127  
Qy 5314 CTACATCGGGCCGATGGCTCCTTGGCGCTGTGCTGCCCAACGGCATGGAGGTGGCGCT 5373

Db 5128 CCAGATTGGTTATGACGGCTCCCTCAGAAATTATCTACGCCAGTGGCTGGACTCACACTA 5187  
Qy 5374 GCAGACTGAGCCCCACTTCTCTGGCTGGCAACCGTCAACCCACCGTGGCAAGAGGAATGT 5433  
Db 5188 CCAAACAGAGCCGCACTTCTGGCTGGCAACCGTCAATCCGACGTTGCCAAAAGAAACAT 5247  
Qy 5434 CACGCTGCCCATCGACRAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAAAGACAGGC 5493  
Db 5248 GACTTTGCCTGGCGAGAACGGTCAAAACTTTGGTGGAAATGGAGATTCGGAAGAGCAAGC 5307  
Qy 5494 TCGGGCCAGGTCACTGTCTTTTGGCGCCCGCTGCGGGTGCAACAACCGAAATCTCCTATC 5553  
Db 5308 CCAAGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAAACCTCCTTTC 5367  
Qy 5554 TCTGGACTTTGATCGCGTAAACACGCACAGAGAAGATCTATGATGACCACCCGCAAGTTTCA 5613  
Db 5368 AGTTGACTTTGATCGAACAAACAAAGACAGAAAAGATCTATGACGACCAACCGTAAATTTCT 5427  
Qy 5614 CTTTCGGATTCTGTACGACAGCGGGGGCGGCGCCAGCCTCTGGTCAACCCAGCAGCAGGCT 5673  
Db 5428 ACTGAGGATCGCCTACGACACGTCTGGGCAACCCGACTCTCTGGCTGCCAAGCAGAGCT 5487  
Qy 5674 GAATGGTGTCAACGTTGACATACTCCCTGGGGGTTCATATTGCTGGCATCCAGAGGGGCAT 5733  
Db 5488 GATGGCCGTCAATGTACCTATTTCATCCACAGGTCAAAATGGCAGCATCCAGCGAGGCAC 5547  
Qy 5734 CATGTCTGAAAGAAATGGAATACAGCACAGCGGGCGCGCATCATCCAGGATCTTCGCTGA 5793  
Db 5548 CACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGCTCTTGTCTGA 5607  
Qy 5794 TGGGAAGACATGGAGCTACACATACTTAGAGAAAGTCCATGGTGTCTGTACTACAGGCCA 5853  
Db 5608 TGCTAAAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTCTCTTCTGCTTCATAGCCA 5667  
Qy 5854 GAGCAGTATATCTTTGAGTTCGACAAAGATGACCGCCTCTCTTCTGTGACGATGCCCAA 5913  
Db 5668 GCGCAGTACATCTTCGAATACGATATGTGGACCGCCTGTCTGCCATCACCATGCCCAG 5727  
Qy 5914 CGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAAACATCTATCA 5973  
Db 5728 TGTGGCTCGCCACACCATGAGACCATCCGATCCGATCCATTTGGTACTACCGCAACATATACA 5787  
Qy 5974 GCCCCTGAGGGCAATGCCTCAGTCAATACAGGACTTCACTGAGGATGGGCACCTCCTTCA 6033  
Db 5788 CCCCCGGAAGCAACGCCCTCCATCATCAGGACTACACGAGGAAGGGCTGCTTCTACA 5847  
Qy 6034 CACCTTCTACCTGGGCACTGGCCGCGAGGTGATATACAAGTATGGCAAACCTGTCAAAGCT 6093  
Db 5848 AACAGCTTTCTTGGGTACAAGTCGAGGGTCTTATTCAAATACAGAAGGCAGACTAGGCT 5907  
Qy 6094 GGCAAGACGCTCTATGACACCACCAAGGTCACTTTCACTTCACTTACCGCTACCGCAGCAT 6153  
Db 5908 CTCAGAAAATTTATATGATAGCAACAAGAGTCAAGTTTACCTATGATGAAAACAGCAGAGT 5967  
Qy 6154 GCTGAAGACCATCAACCTACAGAATGAGGGCTTCACTGCAACCATCCGCTACCGTACAGAT 6213  
Db 5968 CCTAAAGACAGTAAACCTCCAGAGTGTGTTTATTATTCACCATTAGATACAGGCAAT 6027  
Qy 6214 TGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCCCG 6273  
Db 6028 TGCTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGGTAAATGCAAG 6087  
Qy 6274 TTTTGACTACAACTATGACAAACAGCTTCCGGGTGACAGCATGCAAGTGTGATCAACGA 6333  
Db 6088 ATTTGACTATAGCTATGACAAACAGCTTTCGAGTGACAGCATGCAAGGTGTGATCAATGA 6147  
Qy 6334 GACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTT 6393  
Db 6148 AACGCCACTGCCTATTGATCTGTATCAGTTTGTATGACATTTCTGGCAAAGTTGAGCAGTT 6207  
Qy 6394 TGGGAAGTTTGGTGTCACTTTACTATGACATTAACAGATCATCACCACAGCTGTCTATGAC 6453  
Db 6208 TGGAAAGTTTGGAGTTATATATATGATATTAACCAAGATCATTTCTACAGCTGTATGAC 6267

QY	6454	CCACACCAAGCATTTTGTATGCATATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCG	6513
Db	6268	CTATACGAAGCATTTTGTATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTGAG	6327
QY	6514	CTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGCGGGCGAGTAGTGAAGAAGGA	6573
Db	6328	GTCGCTCATGTACTGGATTACAATTGAGTATGATAACATGCGGTGCGGTAAACCAAGAGAGA	6387
QY	6574	GCTGAAGGTAGGACCCCTACGCCAATACCACCTCGCTACTCTATGAGTATGATGCTGACGG	6633
Db	6388	GATTAAATAGGGCCCTTTGCCAACACCACCAAAATATGCTTATGAATATGATGTTGATGG	6447
QY	6634	CCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAA	6693
Db	6448	ACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACCTACGATCTGAA	6507
QY	6694	TGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCCTACGGTATGA	6753
Db	6508	TGGAACCTCCATTTACTGAACCCCAAGTAACAGTGGCGCTGTGACACCCCTTCGGTATGA	6567
QY	6754	CATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATAACAAGATGGATGAGGATGGCTT	6813
Db	6568	CCTGCGAGACAGAATCACTCGACTGGGTGATGTTCAATATAGCTCCAAGGGGCTTCTAACTCGAGT	6627
QY	6814	CCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATCAAGGC	6873
Db	6628	CCTACGTCAAAGGGGCACGGAAATCTTTGAATATAGCTCCAAGGGGCTTCTAACTCGAGT	6687
QY	6874	CTACAAACCGGGCTGGCAGCTGGAGTGTGAGGTACCGCTACGATGGCCCTGGGCGCGCGT	6933
Db	6688	TTACAGTAAAGGCAGTGGCTGCACAGTGATCTACCGTTATGACGGCCTGGGAAGGCGTGT	6747
QY	6934	GTCCAGCAAGAGCAGCCACAGCCACCACCTGCAGTCTTCTATGCAGACCTGACCAACCC	6993
Db	6748	TTCTAGCAAAACCCAGTCTAGCACAGCACCTGCAGTTTCTTATGCTGACTTAACCTATCC	6807
QY	6994	CACCAAGGTCAACCCACCTGTACAAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGA	7053
Db	6808	CACTAGGATTACTCATGTCTCAACCACTTCGAGTTCAGAAATTACCTCCCTGTATTATGA	6867
QY	7054	CTTGCAAGGACACCTCTTGCCATGGAGCTGAGCAGTGGTGATGAGTTTACATAGCTTG	7113
Db	6868	TCTCCAAGGACATCTTTTGCCATGGAAATCAGCAGCTGGGGATGAATTCATATTGTCATC	6927
QY	7114	TGACAAACATCGGGACCCCTCTTGCTGTCTTTAGTGGAAACAGGTTTGATGATCAAGCAAAT	7173
Db	6928	GGATAACACAGGGACACCACTGGCTGTGTTTCAGTAGCAATGGGCTTATGCTGAAAACAGAT	6987
QY	7174	CCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCAACTTTTCAGATCATCAT	7233
Db	6988	TCAGTACACTGCATATGGGAAATCTATTTTGACTCTTAATATTGACTTTTCACTGGTAAT	7047
QY	7234	AGGCTACCATGGTGGCCCTCTATGATCCACTCAACAGCTTGTTCCACATGGCGCGCGAGA	7293
Db	7048	TGGATTTTCATGGTGGCCCTGTATGACCCCACTCACCAAATTAATCCACTTTGGAGAAAAGA	7107
QY	7294	TTATGATGTGTGGCCGACGCTGGACTAGCCCGACCAACGAGCTGTGGAAAGCACTTAG	7353
Db	7108	TTATGACATTTTGGCAGACGGTGGACAACACCTGACATAGAAATCTGGA--AAAGAAAT	7164
QY	7354	TAGCAGCAACGTCAATGCTTTTAAATCTCTATATGTTCAAAAAACAACCCCATCAGCAA	7413
Db	7165	TGGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCCCTGCAAGCAA	7224
QY	7414	CTCCAGGACATCAAGTGCTTTCATGACAGATGTTTAAACAGCTGGCTGCTCACCTTTGGATT	7473
Db	7225	AATCCATGACGTGAAGATTACATCACAGATGTTTAAACAGCTGGCTGGTGACATTTGGTTT	7284
QY	7474	CCAGCTACACAACGTATCCCTGGTTATCCCAAACCAAGACATGGATGCCATGGAAACCTTC	7533
Db	7285	CCATCTGCACAAATGCTATCTCTGGATTCCCTGTTCCCAAATTTGATTTAACAGAAACCTTC	7344

QY	7534	CTACGAGCTCATCCACACACAGATGAAAAACGAGGAGTGGGACAAACAGCAAGTCTATCCT	7599
Db	7345	TTACGA-----ACTGTGAAGAGTCAGCAGTGGATGATATACCGCCCATCTT	7392
QY	7594	CGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTGTCACTTAGAACGGTTTGA	7653
Db	7393	CGGAGTCCAGCAGCAAGTGGCGCGCAGGCCAAGCCCTCCTGTGCTGGGAAGATGGC	7452
QY	7654	CCAGCTCTATGGTCTCCAAATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAAGTTTGC	7713
Db	7453	CGAGGTGCAGGTGAGCCGGCGCCGGCCGGC---GGCGCGCAGTCCTGGCTGTGGTTCCG	7509
QY	7714	ATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAGTTCCTTTGAAGGATGGCCGAGTGAC	7773
Db	7510	CACGGTCAAGTCGCTCATCGGCAAGGGCGTCATGCTGGCCGTACGCCAGGGCCGCTGCA	7569
QY	7774	CACAGACATCATCTGTGTGGCCCAATGAGGATGGGCGAAGGTTGCTGCCATCTTGAACCA	7833
Db	7570	GACCAACGTGTCTAACATCGCCAAACGAGGACTGCATCAAGGTGGCGCCGTCTCAACAA	7629
QY	7834	TGCCCACCTACCTAGAGAACCTGCACCTTCAACCTGATGGGGTGGATACCCATTACTTTGT	7893
Db	7630	CGCCTTCTACCTGAGAAACCTGCACCTTCAACCTGAGGGCAAGGACACGCACTACTTCAT	7689
QY	7894	GAACCCAGGACCTTCAGAAGGTGACCTGGCCATCCTGGCCCTCAGTGGGGCGCGGAAC	7953
Db	7690	CAAGACCACCAACGCCGAGAGCGACCTGGGACGCTGCGGTTGACCAAGCGCCGCAAGGC	7749
QY	7954	CCTGGAGAAATGGGTCAACGTCACTGTGTGCCAGATCAACACAGTACTTAATGGCAGGAC	8013
Db	7750	GCTGGAGAACGGCATCAACGTGACGGTGTGCGAGTCCACCACGGTGGTGAACGGCAGGAC	7809
QY	8014	TAGACGCTACACAGACATCCAGCTCCAGTCCAGTACGGGGCACTGTGCTTGAACACACGCTACGG	8073
Db	7810	GCGCAGGTTCCGGACCTGGAGATGCAAGTTCCGGCGGCTGGCGCTGCACGTGCGCTACGG	7869
QY	8074	GACAAACGTTGGATGAGGAGAAGGCACGGGTCTGGAGCTGGCCCGGCAGAGAGCCGTGCG	8133
Db	7870	CATGACCCTGGACAGGAGAAGGCGCGCATCTGGAGCAGGCGCGGCAGCGCGCTCGC	7929
QY	8134	CCAAGCGTGGCCCGCAGCAGCAGAGACTGGGGGAAGGGAGGAAGGCTGCGGGCCTG	8193
Db	7930	CCGGGCTTGGCGCGCAGCAGCAGCGCGTGGCGCAGCGGCGAGGAGGGCGCGCCTCTG	7989
QY	8194	GACAGAGGGGAGAAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAAGGCTACGACGGCTT	8253
Db	7990	GACGAGGGCGAGAAGCGGCAGCTGCTGAGCGCCGGCAAGGTGCAGGGGTACGACGGGTA	8049
QY	8254	TTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTCAGACAGCGGCCCAACACATCCACTT	8313
Db	8050	CTACGTACTCTCGGTGGAGCAGTACCCCGAGCTGGCCGACAGCGCCCAACACATCCAGTT	8109
QY	8314	CATGAGACAGCGCAGATGGGCCGGAGGTGAC	8345
Db	8110	CCTGCGGACAGCAGATCGGCAGGAGGTAAC	8141

RESULT 9  
US-09-808-602-78  
; Sequence 78, Application US/09808602  
; Patent No. US20020155115A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herrman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: No. US2002015511  
; FILE REFERENCE: 15966-697 CIP  
; CURRENT APPLICATION NUMBER: US/09/09-



Db 1977 ACAGTATTCAAAGGACATGCCAGTGTCTACAGTGGCTGGAAGAGCAGAAATGCCATGT 2036  
QY 1909 GCCCACCACACAGTGTATCGATGTGGCTGCAGCAACCATGGCACCTGCATCACGGGCAC 1968  
Db 2037 GCCCATGAACCAAGTGCATCGATCCCTTCTGTGGGGCCACGGCTCCTGCATGTATGGGAA 2096  
QY 1969 CTGCATCTGCAACCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCC 2028  
Db 2097 CTGCGTGTGTCAGCTGGCTACAAGGGCGAGCACTGCGAAGAAGTGGATTGCTTGGATCC 2156  
QY 2029 CACATGTTTCAGSCGGGGTGTCTGCTGAGAGGCGGAATGCCATTGCTTTGTGGGATGGG 2088  
Db 2157 AACCTGCTCCAGCCATGGTGTCTGTGTGAACGGAGAGTGTCTATGCAGCCCCGGCTGGG 2216  
QY 2089 AGGCACCAACTGCGAGACCCCCAGGGCCACATGCTTAGACCAAGTTCAGGCCACCGAAC 2148  
Db 2217 CGGGCTCAACTCGAGCTGGCGAGGTCCAGTGCCAGACCACTGTAGTGGCATGGCAC 2276  
QY 2149 CTTCTCCCGSACACCGGGCTTTGCACTGTGACCCAAAGCTGGAAGTGGACTGGACACGACTGTC 2208  
Db 2277 TTACCTCCCTGACTCTGGCCTCTGCAACTGTGATCCGAATTGGATGGTCCCGACTGCTC 2336  
QY 2209 TATCGAGATCTGTGCTGCCGACTGTGTGGCCCATGGCGTGTGCGTAGGGGGCACCTGCCG 2268  
Db 2337 TGTGAAGTGTCTCAGTAGACTGTGGCACTCACGGCGTCTGCACTCGGGGAGCCTGCCG 2396  
QY 2269 CTGCGAGGATGCTGTGATGGGGGAGCTGCGACCGAGCGGGCCCTGCCACCCCGCTGTGC 2328  
Db 2397 CTGTGAAGAGGGCTGGACAGGGCGGGCTGTGACCAAGCGGTGTGCCACCCCGCTGCAT 2456  
QY 2329 CGAGCATGGGACCTGCCGGCAGCGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACA 2388  
Db 2457 TGAGCACGGGACCTGTAAAGATGGCAATGTGAATGCCGAGAGGGCTGGAATGGTGAA 2516  
QY 2389 CTGCACCATCGCTCACTATCTGGATAGGCTAGTTAAAGAGGGTTCCTGGCTGTGCA 2448  
Db 2517 CTGCACCATGAT-----GGCTGCCCTGATTTGTGCAA 2549  
QY 2449 TGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTG 2508  
Db 2550 CGGTAACGGGAGATGCACACTGGGTGCAACAGCTGGCAGTGTCTGCCAGACCGGCTG 2609  
QY 2509 GAGAGGAGCTGGTGTGACACTTCCATGGAGACTGCCCTGCGGTGACAGCAAGACAATGA 2568  
Db 2610 GAGAGGGCCCGATGCAACGTTGCCATGGAAACCTCTCGCTGTATGAACAAGGATAATGA 2669  
QY 2569 TGGAGATGGCTGGTGGACTGCATGGACCTGACTGTGCTGCCCTCAGCCCTGTGCCATAT 2628  
Db 2670 GGGAGATGGCTGGTGGACTGCCCTGACCTGACTGTGCTGCCCTCAGTCCAGCTGTGCA 2729  
QY 2629 CAACCCGCTGTGCTTGGCTGCCCTAACCTCTGGACATCATCAGGAGACACAGGTCCC 2688  
Db 2730 CAGCCTGCTCTGTGCGGGGTCTCGGGACCCCTTGGACATCATTCAGCAAGGCCAGACAGA 2789  
QY 2689 TGTGTACAGCAGAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGCAGGGA 2748  
Db 2790 CTGGCCTGCGGTGA-----AGTCCCTTATGATCGTATCAAGTCTTGGCAGGCAAGGA 2843  
QY 2749 CAGCAGGCACATAATCCCCGGGAGAACCCCTTTGATGGAGGSGATGCTTGTGTTATTCG 2808  
Db 2844 CAGCACCCACATCATCTCTGGAGACACCCCTTCAATAGCAGCCTGGTGTCTCTGATCCG 2903  
QY 2809 TGGCCAAAGTGTACATCAGATGGAACCCCTTGGTGGTGTGAACATCAGTTTTGTCAA 2868  
Db 2904 AGGCCAAAGTAGTAACCAACGATGGGACCCCTTGGTGGTGTGAATGTGTCTTTGTCAA 2963  
QY 2869 TAACCTCTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTTGGTGACAAA 2928  
Db 2964 GTACCCAAAATATGGCTACACCATCACTGCCAGGACGGCAGCTTTGACCTGATGGCAA 3023  
QY 2929 TGGCGGCATCTCCATCATCTCGGGTTGAGCGGGCACTTTTCATCACACAGGAGCACAC 2988

Db 3024 TGGGGGCTCTGCCTTGACTCTTCACTTTGAGCGAGCCCTTTTCATGAGCCGGGAGCGCAC 3083  
QY 2989 CCTGTGGCTGCCATGGGATCGCTTCTTTGTGTATGGAACCATCATCATGAGACATGAGA 3048  
Db 3084 AGTATGGCGCGCGTGAACAGCTTCTATGCCATGGACACCCCTGGTAAATGAAGACGGAGGA 3143  
QY 3049 GAATGAGATTCACAGCTGTGACCTGAGCAATTTTGCCCGCCCAACCCAGTCTCTCTCC 3108  
Db 3144 GAACTCCATCCCAGCTGTGACCTCAGTGGCTTTGTCCGSCCTGATCCGATCATCATCTC 3203  
QY 3109 ATCCCCACTGACGCTCCTTCGCCAGCTCTGTGCAGAGAAAGGCCCCCATTTGTCCCGGAAAT 3168  
Db 3204 CTCTCCTCTGTCCACCTTCTTCAGCGCTTCCCCTGCGGCGAACCCTCATTTGTGCTGAGAC 3263  
QY 3169 TCAGGCTTTGCAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAG 3228  
Db 3264 CCAGGTTCTTCATGAGGAGATCGAGCTCCCTGGCACCAACGTGAAGCTCCGTTACCTCAG 3323  
QY 3229 CAGCCGGACCCCTGGCTACAAATCTGTCTCTGAGGATCAGCCTCACCCACCCGACCATCCC 3288  
Db 3324 CTCAGAACACAGCAGGGTACAAGTCACTGCTGAAGATCACCATGACCCAGTCCACGGTGCC 3383  
QY 3289 CTTCAACCTCATGAAGGTGCACCTCATGGTAGCGGTGGAGGCCGCCCTCTTTCAGGAAGTG 3348  
Db 3384 CTTGAACCTCATCCGGGTTCACTTGAATGGTTGCCGTGGAGGGCATCTCTTCCAGAAAGTC 3443  
QY 3349 GTTCGCTGCAGCCCCCAGACCTGTCTTATTTTTCATTTGGGACAAGACAGACGTCTACAA 3408  
Db 3444 GTTCCAGGCTTCTCCAAACCTGGCTACACATTCATCTGGGACAAGACAGACGCTTATGG 3503  
QY 3409 CCAGAAGGTGTTGGGCTTTTCAGAAGCCCTTTGTTCCTGGGTATGAATATGAATCTCTG 3468  
Db 3504 CCAAAGGTTTATGGCCTATCGGATGCTGTTGTGTCTGTGGATTTGAATATGAGACCTG 3563  
QY 3469 CCCAGATCTAATCCTGTGGGAAAAAGAACAAACAGTGTGAGGGCTATGAAATGAGCGC 3528  
Db 3564 CCCAGTCTCATCTGTGGGAAAAAGAACAGACAGCCCTACTTCAAGGATTCGAGCTGGACCC 3623  
QY 3529 GTCCAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTCAACATTCAAAGTGGTAT 3588  
Db 3624 TTCCAACCTTGGTGGCTGGTCCCTGGATAAGCACACACCTCAATGTGAAAAGCGGAAT 3683  
QY 3589 CCTGCACAAAGGGAATGGGAGAACCCAGTTTGTGTCTCAGCAGCCTCCTGTCAITGGGAG 3648  
Db 3684 ACTACTCAAAGGCACAGGGGAGAACCCAGTTCCTGACCCAGCAGCCCGCCATCATCACCAG 3743  
QY 3649 CATCATGGGCAATGGCGCCGAGAGAGCATCTCCTGCCCCAGCTCAACGGCCCTTGCTGA 3708  
Db 3744 CATCATGGGTAAACGGTCCGCCGAGAGGATCTCCTGTCCCAGTGAATGGCCCTTGCTGA 3803  
QY 3709 CGGCAACAAGCTCCTGGCCCCCAGTGGCCCCCTCACCTGTGGCTCTGACGGGAGCCCTCTATGT 3768  
Db 3804 AGGCAACAAGTGTGGCCCCCGTGGCCCCCTGGCTGTGGGATCGATGGGAGCCCTCTTTGT 3863  
QY 3769 GGGTGATTTCAACTACATTAGAAGGATCTTCCCCTCTGGAATGTCAACCAACATCCTAGA 3828  
Db 3864 CGGTGACTTCAATTATATCCGGCGCATCTTCCCCTCTCGAAACGTGACCAAGTATCTTGA 3923  
QY 3829 GCTGAGGAATAAAGATTTTCAGACATAGTACAGTCCAGCACACAAATACTACCTGGCCAC 3888  
Db 3924 GTTACGAAATAAAGAGTTTAAACATAGCAACAGCCCGAGGACACAAGTACTACTTGGCTGT 3983  
QY 3889 AGACCCCATGAGTGGGGCCGTCTTCTCTTGACAGCAACAGCCGGCGGCTCTTTAAAT 3948  
Db 3984 GGACCCCTGTGACTGGCTCGCTCTATGTCTCTGACACCAACAGTCCCGGATCTACCCAGT 4043  
QY 3949 CAAGTCCACTGTGGTGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGG 4008  
Db 4044 CAAGTCTCTAAGCGGAGCCAAAGACCTGGCTGGGAATTCGGAAGTGTGCGCCGGGACTGG 4103  
QY 4009 TGACCAAGTGCCTCCCTTTTGTATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAGC 4068  
Db 4104 CGAACAAATGTCTACCCCTTTGTATGAAGCCCGCTGTGGGATGGCGGGAAGGCTGTGGATGC 4163



Db 6321 GGACAAGCAGATTACAGGTTCTCTGAGGAAGGCATGATCAACGCCAGGTTTCGATTACAC 6380  
QY 6285 --ACTATGACACAGACTTCCGGGTGACGAGCATGCAGGCTGTGATCAACGAGACCCCACT 6342  
Db 6381 CTACCACGACACAGCTTCGGCATCGCGAGCATCAAGCCCGTTCATCAGTGAGACTCCCT 6440  
QY 6343 GCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTT 6402  
Db 6441 TCCCGTTGACCTCTACCGCTACGATGAGATTTCTGGCAAGGTGGAACACTTCGGCAAGTT 6500  
QY 6403 TGGTGTCTATTTACTATGACATTAACCAAGATCATCACCACAGCTGTCTATGACCCACACCAA 6462  
Db 6501 CGGGGTCTACTACGACATCAACCAAGATCATCACCACCTGCCGTCTATGACACTCAGCAA 6560  
QY 6463 GCATTTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCGGCTCGCTCAT 6522  
Db 6561 GCATTTTGACACCCATGGGCGCATCAAGGAAGTGCAGTATGAGATGTTCCGGTCCCTCAT 6620  
QY 6523 GTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTAGTGAAGAGGAGCTGAAGGT 6582  
Db 6621 GTACTGGATGACCGTGCAATATGACAGTATGGGCGAGGTCTATCAAGAGGGAACCTGAAACT 6680  
QY 6583 AGGACCTTACGCCAATACCACTCGCTACTCTCTATGATGATGATGATGATGATGATGATGAT 6642  
Db 6681 GGGGCCCTATGCCAACACCAAGTACACCTATGACTACGACGGGAGCGGCCAGCTCCA 6740  
QY 6643 GACAGTCTCCATCAATGACAAGCCACTCTCTGGCGCTACAGCTACGACCTCAATGGGAACCT 6702  
Db 6741 GAGTGTGGCGCTCAATGACCGGCTACCTGGCGTTATAGCTATGACCTCAATGGGAACCT 6800  
QY 6703 GCATTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGGGA 6762  
Db 6801 GCACCTGTAAACCCAGGAACAGTGTCCGCTCATGTCGCTTACGCTATGACCTCCGTGA 6860  
QY 6763 CCGCATCACTCGGCTGGGTGACGTGCAATACAAAGATGGATGAGGATGGCTTCCTGAGGCA 6822  
Db 6861 CCGGATAACCAAGCTAGGGACGTGCACTACAAATCGATGATGATGGCTATTTATGCCA 6920  
QY 6823 GCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCTGCTCATCAAGGCTTACACCG 6882  
Db 6921 GAGAGGATCTGACATCTTTGAATACAACTCCAAGGCTTCTAACGAGAGCGTACACAA 6980  
QY 6883 GGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCTGGCGCGCGCGCTGTCCAGCAA 6942  
Db 6981 GGCCAGCGGTGGAGTGTGCAGTACCGCTATGATGGCGTGAGCGCGCGGCTTCCTACAA 7040  
QY 6943 GAGCAGCCACAGCCACCACTGCACTTCTATGACAGACCTGACCAACCCCAACCAAGGT 7002  
Db 7041 GACCAACCTGGGCCACCACTACAGTACTTCTATTCGACCTCCACCAACCAACGAT 7100  
QY 7003 CACCCACCTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGAAGG 7062  
Db 7101 CACCCATGTTTACAACCACTCCAACCTCTGAGATCACTCTACTACTATGACCTCCAGG 7160  
QY 7063 ACACCTTTTGCCATGGAGCTGAGAGTGGTGTAGTGTGATGATGATGATGATGATGATGAT 7122  
Db 7161 CCACCTCTTTGCCATGGAGAGCAGTAGTGGGGAAGAGTACTATGTTGCTTCAGATAACAC 7220  
QY 7123 CGGACCCCTCTTGTCTGTTTGTAGTGAACAGGTTTGTATGATCAAGCAAACTCCTGTACAC 7182  
Db 7221 CGGACTCTCTGGCTGTTTGTAGTATCAATGGCCTCATGATCAAGCAAACTCCAATACAC 7280  
QY 7183 AGCCTATGGGAGATCTACATGGATACCAACCCCACTTTTCTGATCATCATAGGCTACCA 7242  
Db 7281 AGCCTATGGGAGATTTACTATGACTCCAATCCAGACTTTTCTGATGATGATGATGATGATGAT 7340  
QY 7243 TGGTGGCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGGCGAGATTATGATGT 7302  
Db 7341 CGGAGGCTCTACGACCCCTCACCAAGCTCGTTTCACTTTTACGACGCTGATTATGACGT 7400  
QY 7303 GCTGGCCGAGCTGGACTAGCCACAGACCAAGCTGTGGAAGCACCTTAGTAGCAGCAA 7362

Db 7401 GCTGGCAGGACGGTGGACGTCTCCCGGACTACACCATGTGGAGAA---TGTGGGCAAGGA 7457  
QY 7363 CGTCATGCCCTTTAATCTCTATATGTTTCAAAAACAACACCCCATCAGCAACTCCCAGGA 7422  
Db 7458 GCCAGCCCTTCAACCTGTACATGTTCAAGAAACAACATCCACTCAGTAATGAGCTGGA 7517  
QY 7423 CATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACA 7482  
Db 7518 TTTAAAGAACTACGTGACAGACGTGAAGAGCTGGCTCGTGTGTTTGGATTTCAGCTCAG 7577  
QY 7483 CAACGTGATCCCTGTTATCCCAAACCAAGAGTGGGCAACAGCAAGTCTATCTCGGGGTACA 7542  
Db 7578 CAACATCATCTCTGATTCCCAAGAGCCAAAATGTATTTTGTGCTCCCCCTATGA--- 7634  
QY 7543 CATCCACACACAGATGAAAACGCGAGAGTGGGCAACAGCAAGTCTATCTCGGGGTACA 7602  
Db 7635 -----ACTGTCAAGAGAGCCAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7685  
QY 7603 GTGTGAAGTACAGAAGCAGCTCAAGGCTTTGTCACTTTAGAACGGTTTGACCAAGTCTA 7662  
Db 7686 GCAGACAACAGAGAGGCTAACCAGGCTTTCTGGCTCTAGAAGGACAGGTCATCTCTAA 7745  
QY 7663 TGGCTCCACAATCACAGCTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGATCCAGCGG 7722  
Db 7746 AAAGCTCCATGCAGGCATCCGAGAGAAAGC-----AGGCCACTGGTTTGTACGACCAC 7799  
QY 7723 CTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTTGAAGGATGGCCGAGTGACCAAGACAT 7782  
Db 7800 GCCCATCATCGGCAAGGCATCATGTCGCCATCAAAGAAAGGCGGTGACCAAGCGGT 7859  
QY 7783 CATCAGTGTGGCCAATGAGGATGGCGGAAGGTTGCTGCCATCTTGAACCATGCCCACTA 7842  
Db 7860 GTCTAGCATCGCCAGTAGGACAGCCGCAAGGTAGCATCCGTGTTGAACAACGCTCTA 7919  
QY 7843 CCTAGAGAACCTGCACCTTCAACATGATGGGTGGATACCCATTACTTTGTGAACCCAGG 7902  
Db 7920 CTTGGACAAGATGCACACTACAGCATCGAGGGCAAGGACACACACTACTTCTGTAAGATCG 7979  
QY 7903 ACCTTCAAGAGTGACCTGGCCATCTCTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAA 7962  
Db 7980 TGCAGCGGACGGTGACCTGGTTACGCTGGGGACCAACCATTTGGCGCAAGGTGCTGGAGAG 8039  
QY 7963 TGGGTCAACGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTA 8022  
Db 8040 CGGGGTGAACGTGACCGTGTACAGCCCAAGTGTGTTGTTGAACGCGCAGGACTCGAAGGT 8099  
QY 8023 CACAGACATCCAGTCCAGTACGGGCACTGTGCTTGAACACACGCTACGG----- 8073  
Db 8100 CACCAACATGAAATTCAGTACTCCAGTACTCCAGCTGTGCTCAGCATACGCTACGGCTCACCCC 8159  
QY 8074 GACAACTGTTGATGAGGAGAAGGACGCGGTCTCTGGAGCTGGCCCGCAGAGAGCCGTGG 8133  
Db 8160 CGACACACTGGATGAAGAGAAGGCGCGCTCTCTGGACCAAGCGCGACAGAGGCGCTGG 8219  
QY 8134 CCAAGCTGGGCGCGGAGCAGCAGAGACTGCGGGAAGGGAGGAAGGCTGCGGGCGCTG 8193  
Db 8220 TACTGCTGGCCAAAGGAGCAGCAGCAAGAAAGCCAGGACGCGGAGAGAGGCGCTCTGTG 8279  
QY 8194 GACAGAGGGGAGAAAGCAGCAGGTCGTGAGCACAGGCGGCTGCAAGGCTACGACGCTT 8253  
Db 8280 GACGAGGGGAGAAAGCAGCAACTCTGAGCAAGGACGCGGTCAAGGTTATGAGGGCTA 8339  
QY 8254 TTTCTGTGATCTCTGTGAGCAGTACCCAGAACTGTGACAGAGCGCCCAACAACATCCACT 8313  
Db 8340 TTACGTGCTTCCGTTGGAACAGTACCCAGAGCTGGCAGAGTAGCAGCAACATCCAGTT 8399  
QY 8314 CATGAGACAGAGCGAGATGGGCGGAGGTGACAGA 8348  
Db 8400 CTTAAGACAGAAATGAGATGGGAAAGAGGTAACAA 8434

; Sequence 66, Application US/09900198  
; Publication No. US20030087816A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Cornie AM  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herrmann, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishna  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-697  
; CURRENT APPLICATION NUMBER: US/09/800,198  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 66  
; LENGTH: 8689  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-09-800-198-66

Query Match 40.7%; Score 3397.8; DB 10; Length 8689;  
Best Local Similarity 64.8%; Pred. No. 0;  
Matches 5469; Conservative 0; Mismatches 2712; Indels 254; Gaps 19;

QY	35	ATGGACGTGAAGGAGAGAAACCTTACCGCTCGCTGACCCCGCGCGGACGCGCGAGCGC	94
Db	133	ATGGATGTGAAGGATCGGCGACATCGCTCTTTGACCAGGGGACGGTGTGGCAAGGAGTGT	192
QY	95	CGCTACACGAGCTCGTCGGGACAGCGAGGAGGCAAAAG---CCCCGAGAAATCGTAC	151
Db	193	CGCTACACGAGCTCCTCTGTGACAGTGAGGACTGCGGTGTGCCACCGAGAACTCTAC	252
QY	152	AGCTCCAGCGAGACCTTGAAGCCCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCGC	211
Db	253	AGTTCCAGTGAGACCCCTGAAGCCTTATGACCATGACAGCAGATGCACATATGGAACCGA	312
QY	212	GTCAAGGACATTGTGCCGCGAGGAGCGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCAC	271
Db	313	GTACAGACCTGGTGCACCGGAGTCCGATGAGTTTCTAGACAAGGGGCTAATTCACC	372
QY	272	CTGCGGGAGCTGGGGCTGGAAGAAGTAAACGCCCCCTCACGGGACCCCTGTACCGGACAGAC	331
Db	373	CTGGCAGAAATTGGGAATCTGCGAG---CCCTCCCCACACCGAAGTGGTTACTGTTCCGAC	429
QY	332	ATTGGCCTGCCCAATCGCGCTACTCCATGGGGCTGGCTCTGTATGCCACATGGAGGCT	391
Db	430	ATGGGGATCCTCCACAGGGCTACTCCCTGAGCACTGGGTCTGTATCGGACTCGGACACC	489
QY	392	GACACGGTGTGTCCCCTGAGCACCCCGTGGTGTGGGGCCGGAGCAGCAGCGTCAAGG	451
Db	490	GAGGAGGGATGTCTCCAGAAACATGCCATCAGACTGTGGGACGAGGGATAAAATCGAGG	549
QY	452	CGCAGCTCCTGCTGTCCAGCCCGGCCAAATTCOAATCTCACACTCACCGACACCGAGCAT	511
Db	550	CGCAGCTCTGGCTTGTCCAGCCGCGAGAACTCAGCCCTTACTCTGACTGATTCTGACAAT	609
QY	512	GAAAAACACTGAGACTGATCATCCGGCGGCC---TGCAGAACCCACGCGGGCTCCGGAC	567
Db	610	GAAAAATAAATCGGATGACGACAATGGTCGACCCATTCACCTACATCTCTGCTAGCCTC	669
QY	568	GCCGCGCGCGCGCTCTCGCACGCCACACCCCCAAACAGCACACCGCGCCTCCATTAA	627
Db	670	CTCCCATCTGCTCAGCTGCCTAGCTCCCATATCTCTCCACCAAGTTAGTGCC-AGATGCC	728
QY	628	CTCCCTGAACCGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGG-----	674
Db	729	ATTGCTAGACAGCAACACCTCCCATCAGATCATGAGATCATGACACCAACCCCGATGAGGAATTCTC	788

QY	675	-----CCCCACGACCACTCGCTCTCCGGAGAGCCCCCTGCCGGCGGCGCCC	722
Db	789	CCCTAATTATACCTGTCTAGAGCATGTCTCAGGGCCCCAGCAAGCCTCCAGTAGTGGCCC	848
QY	723	AGGAGCCTGCCACGCCCCAGAGAACTGGCTGTCTCAACAGCAACATCCCCCTGGAGACCA	782
Db	849	TCCGAACCAACACAGCCAGTCAACGCTGAGGCCCTCTGCCACCTCCTCATAAACACAC	908
QY	783	GGAACCTAGGAAGCAGCCATTCTCTAGGACATTGCAAGCAACCTCATTTGAGATGGACA	842
Db	909	CCTGTCCCACCACTCCTCTGCAACTCCCTCAACAGAAACTCACTGACCAATCGCGG	968
QY	843	TTCTCGGCGCCTCCCGCCA-----	861
Db	969	GAGTCAAAATCCACGCCCCAGTCTCTGCAACCAATGACCTGGCCACCAACGCGGAGTCCGT	1028
QY	862	-----TGATGSGGCTTACAGTGACGGGCATTTCT	891
Db	1029	TCAGTCCAGGACAGCTGGTGTGAACAGTAAACGTGCGCTGGAGACGCGGCACTTCT	1088
QY	892	CTTCAA---GCCTGGAGGACCTTCCCCGCTCTTCTGCACCACATCACCCAGGGTATCCACT	948
Db	1089	CITCAAGACGTCCTCCGGAAGCACACCCCTGTTTCCAGAGCTCTTCTCCAGGATACCCCTT	1148
QY	949	GACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCCCGCAGCACCTTCGCCCCG	1008
Db	1149	GACCTCAGGACCGTTTATACACCAACCAACCCCGCTGCTGCCACGGAATACATTTCTTAG	1208
QY	1009	GCCGCTTTAACTCAAGAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAG	1068
Db	1209	GAAAGCCTTCAAGCTGAAGAAACCCCTCCAATACTCGAGTTGAAATCGCGCCGCTGTCTC	1268
QY	1069	CGCATCGTCATCTCAGCCACTCTGGTTCATCTCTGCTGGCATACTTTGTGGCCATGCACCT	1128
Db	1269	TGCCATGCGCGTGCCTCTCTGGCCATTTTGTGGCCTATTTTCATAGCAATGCATCT	1328
QY	1129	GTTTGGCCTAAACTGGACCTGCGACCCGATGGAGGGCAGATGTATGAGATCACGGAGGA	1188
Db	1329	GCTCGGACTCAATTGGCAACTCCAGCCGCGAGATGGACACACCTTTAACA-----ATGG	1382
QY	1189	CACAGCCAGCAGTTGGCCTGTGCCAACCCGACGTCTCCCTATATACCCCTCAGGGGGCACTGG	1248
Db	1383	CGTAAGGACCGGCTTACCAGGAAACGATGATGTGGCAACAGTGCCATCTGGAGGCAAGT	1442
QY	1249	CTTAGAGACCCCTGACAGGAAAGGCAAGGAACCAAGAAAGGAAAGCCCACTAGTTTCTT	1308
Db	1443	GCCCTGGTCTGAAAA-----	1459
QY	1309	TCCAGAGGACAGTTTCATAGATTCTGGAGAAATTGATGTGGAAAGGCGGACCTCCAGAA	1368
Db	1460	-----ACAGCAGCATAGACAGCGCGGAGGAGTGGTCGACGGGTGACACAGGA	1511
QY	1369	GATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAA	1428
Db	1512	AGTCCCACGAGGGTGTTTGGAGGTCCCAGATTTCACATCAGTCAGCCTCAGTTCTTAA	1571
QY	1429	ATTCAATGTGTCTCTGGGAAAGGACGCCCTGGTTGGCATTTATGGCAGAAAGGCGCTCCC	1488
Db	1572	GTTCAACATCTCCCTGGGAAAGGATGCCCTCTTCGSGCTCTACATAAGAAAGAGGACTGCC	1631
QY	1489	TCCTTCACATACACAGTTTGACTTTGTGGAGCTGTCTGGATGGCAGGAGGCTCTTAACCCA	1548
Db	1632	ACCATCTCATGCACAGTATGACTTCATGGAAACGCCCTGGACGGAA-----A	1676
QY	1549	GGAGGCGCGAGCCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTGCCCCCTCCAG	1608
Db	1677	GGAGAAGTGGAGTGTGTGCTGAGTCAACCGGAAACCGCGGAGCATCCAGACCTGGTGCA	1736
QY	1609	CCATGAGACAGGCTTCATCCAGTATTTGGATTCAGAAATCTGGCACTTGGCTTTTACAA	1668
Db	1737	GACGAGGCTGTGTTCTCGTGAGTACTTGGATGTGGGCTGTGGCACCTCGCCTTCTACAA	1796
QY	1669	TGACGGAAAGGAGTCAAGAAGTGGTTTCTTTCTCACCACTGCCATTGAGTCGGTGGATAA	1728

Db 1797 TGACGGCAAGGACAAGAGATGGTCTCCTTCAATACGGTTGTCTTAGATTTCAGTGCAGGA 1856

Qy 1729 CTGCCCCAGCAACTGCTATGGCAATGGTGACTGCATCTCTGGGACCTGCCACTGCTTCCT 1788

Db 1857 CTGTCCACGAAACTGCCACGGGAACGGGAATGCGTGTCTGGAAGTGTCTCACTGTTCCC 1916

Qy 1789 GGGTTTCCTGGGCCCCGACTGTGGGACAGACCTCCTGCCCCCGTCTCTGTAGCGGAAATGG 1848

Db 1917 AGGATTCCTAGTGCAGACTGGGCTAAAGCTGCCCTGCCCTGTCTGTGCAGTGGGAATGG 1976

Qy 1849 CCAATACATGAAAGGCAGATGCTTGTGCCACAGTGGCTGGAAGCGCGCTGAGTGCATGT 1908

Db 1977 ACAGTATTCCAAAGGGACATGCCAGTGTCTACAGTGGCTGGAAGAGAGCAGAAATGCCATGT 2036

Qy 1909 GCCCACCACCAAGTATCGATGTGGCTGCAGCAACCATGGCACCTGCATCACGGGCAC 1968

Db 2037 GCCCATGAACCAAGTGCATCGATCCTTCCTGTGGGGCCACCGGCTCCTGCATTTGATGGGAA 2096

Qy 1969 CTGCATCTGCAACCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCC 2028

Db 2097 CTGCGTGTGTGAGCTGGCTACAAAGGGGAGCACTGCGAAGAGTGGATTGCTTGGATCC 2156

Qy 2029 CACATGTTACGGCCGGGTGTCTGCTGAGAGGCGAAATGCCAATGCTTTGTGGATGGGG 2088

Db 2157 AACCTGCTCCAGCCATGGTGTCTGTGTAACGGAGAGTGTCTATGCAGCCCCGGCTGGGG 2216

Qy 2089 AGGCACCACCTGCGAGACCCCCAGGGCCACATGCTTAGACCAGTGTTCAGGCCACCGGAAC 2148

Db 2217 CGGGCTCAACTCGGAGCTGGCGAGGGTCCAGTGCCACAGACCAAGTGTAGTGGCATGGCAC 2276

Qy 2149 CTTCTCCCGGACACCGGGCTTTGCAGCTGTGACCCAGCTGGAAGTGGACTGGACACGACTGTC 2208

Db 2277 TTACCTCCCTGACTCTGGCCTCTGCAACTGTGATCCGAATTGGATGGTCCCGACTGCTC 2336

Qy 2209 TATCGAGATCTGTGCTGCCGACTGTGCTGGCCATGGCGTGTGCGTAGGGGACCTTGCCG 2268

Db 2337 TGTGAAGTGTCTCAGTAGACTGTGGCACTCACGGCGTCTGCATCGGGGAGCCTTGCCG 2396

Qy 2269 CTGCGAGGATGGCTGGATGGGGGACGCTGCGACCAAGCGGGCTGCCACCCGCGCTGTGC 2328

Db 2397 CTGTGAAGAGGGCTGGACAGGCGCGGCTGTGTGACCAAGCGGCTGTGCCACCCCGCTGCAT 2456

Qy 2329 CGAGCATGGGACCTGCCGCGACGGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACA 2388

Db 2457 TGAGCACGGGACCTGTAAAGATGGCAAAATGTGAATGCCGAGAGGGCTGGAATGGTGAACA 2516

Qy 2389 CTGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAA 2448

Db 2517 CTGCACCATGAT-----GGCTGCCCTGATTTGTGCAA 2549

Qy 2449 TGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTG 2508

Db 2550 CGGTAAACGGGAGATGCACACTGGGTGAGAAACAGCTGGCAGTGTGTCTGCCAGACCGGCTG 2609

Qy 2509 GAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCCCTGCGGTGACAGCAAAAGACAATGA 2568

Db 2610 GAGAGGGCCCGGATGCAACGTTGCCATGGAAACCTCCTGCGCTGATAACAAGGATAATGA 2669

Qy 2569 TGGAGATGGCTGGTGGACTGCATGGACCCCTGACTGTGCTGCCAGCCCTGTGCCATAT 2628

Db 2670 GGGAGATGGCTGGTGGACTGCCTGGACCCCTGACTGTGCTGCCCTCAGTCCAGCTGTGAGAA 2729

Qy 2629 CAACCCGCTGTGCCTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGCCCTGAGGATCC 2688

Db 2730 CAGCCTGCTGTGCGGGGTCTCGGGGACCCCTTGGACATCATTCAGCAAGGCCAGACAGA 2789

Qy 2689 TGTGTACAGCAGCAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGCGAGGA 2748

Db 2790 CTGGCCTGCGGTGA-----AGTCCCTCTATGATCGTATCAAGCTCTTGGCAGGCAAGGA 2843

Qy 2749 CAGCAGGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGCGATGCTTGTGTATTCG 2808

Db 2844 CAGCACCCACATCATTCCTGGAGACAAACCCCTTCAATAGCAGCCTGGTGTCTCTGATCCG 2903

Qy 2809 TGGCCAAAGTGATGACATCAGATGGAACCCCTCCTGGTTGGTGTGTAACATCAGTTTTGTCAA 2868

Db 2904 AGGCCAAAGTAGTAACCAACCGATGGGACCCCTCCTGGTGGGTGTGAATGTGTCTTTTGTCAA 2963

Qy 2869 TAAACCTCTCTTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTTGACTTGGTGACAA 2928

Db 2964 GTACCCAAAATATGGCTACACCATCACTCGCAGGACGGCACGTTTGACCTGATTGCCAA 3023

Qy 2929 TGGCGGCATCTCCATCATCCTGCGGTTTCGAGCGGGCACCTTTTCATCACACAGGAGCACAC 2988

Db 3024 TGGGGGCTCTGCCCTTGACTCTTCACTTTGAGCGAGCCCTTTTCATGAGCCGGGAGCGCAC 3083

Qy 2989 CCTGTGGCTGCCATGGATCGCTTCTTTGTTCATGGAACCATCATCATGAGACATGAGGA 3048

Db 3084 AGTATGGCCGCGGTGAACAGCTTCTATGCCATGGACACCCCTGGTGAATGAAGACGGAGGA 3143

Qy 3049 GAATGAGATTCCCAGCTGTGACCTGAGCAATTTTGTCCCGCCCCAACCCAGTCTCTCTCC 3108

Db 3144 GAATCCATCCCAGCTGTGACCTCAGTGGCTTTGTCCGCGCTGATCCGATCATCATCTC 3203

Qy 3109 ATCCCCACTGACGCTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCCATTTGTGCCGGAAT 3168

Db 3204 CTCCTCTGTCCACCTTCTTCAGCGCTTCCCTCGCGGAAACCCCATTTGTGCTGAGAC 3263

Qy 3169 TCAGGCTTTCAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAG 3228

Db 3264 CCAGGTTCTTCATGAGGAGATCGAGCTCCCTGGCACCAACCTGAAGCTCCGTTACCTCAG 3323

Qy 3229 CAGCCGGACCCCTGGCTACAAATCTCTCCTGAGGATCAGCCTCACCCACCCGACCATCCC 3288

Db 3324 CTCCAGAACAGCAGGGTACAAGTCACTGCTGAAGATCACCATGACCCAGTCCACGCTGCC 3383

Qy 3289 CTTCAACCTCATGAAGGTGCACCTCATGCTGAGCGGTGGAGGCGCCCTCTTCAGGAAAGTG 3348

Db 3384 CTTGAACCTCATCCGGGTTCACTTCATGATGGTTGCCGTGGAGGGCATCTCTTCCAGAAAGTC 3443

Qy 3349 GTTCGCTGCAGCCCCAGACCTGTCTCTATTTTCAATTTGGGACAAGACAGACGCTCTACAA 3408

Db 3444 GTTCCAGGCTTCTCCAAACCTGGCTACACATTCATCTGGACAAGACAGACGCTTATGG 3503

Qy 3409 CCAGAAAGTGTGGCTTTTCAGAAAGCCTTTGTTCCTGGGTTATGAATATGAATCCTG 3468

Db 3504 CCAAAGGTTTATGGCCTATCGGATGCTGTGTCTGTGTTGGATTTGAATATGAGACCTG 3563

Qy 3469 CCCAGATCTAATCCTGTGGGAAAAAGAAACAACAGTGTGAGGGCTATGAAATGACGC 3528

Db 3564 CCCAGTCTCATCCTGTGGGAAAAAGGACAGCCCTACTTCAAGGATTCGAGCTGGACCC 3623

Qy 3529 GTCCAAGCTTGGAGGATGGAGCCCTAGACAAACATCATGCCCCTCAACATTCAAAAGTGTAT 3588

Db 3624 TTCCAACCTTGGTGGTGGTCCCTGGATAAGCACACACCCCTCAATGTGAAAAGCGGAAAT 3683

Qy 3589 CCTGCACAAAAGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAGCCTCCTGTCTATTGGGAG 3648

Db 3684 ACTACTCAAAGGCACAGGGGAGAACCAAGTTCCTGACCCAGCAGCCCGCCATCATCACCAG 3743

Qy 3649 CATCATGGGCAATGGGCGCCGGAGAAAGCATCTCCTGCCCCCAGCTGCAACGGCCCTTGTGA 3708

Db 3744 CATCATGGGTAAACGGTCGCCCGCAGAAAGCATCTCCTGTCCAGCTGCAATGGCCCTGTGA 3803

Qy 3709 CGGCAACAAAGCTCCTGGCCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCCTCTATGT 3768

Db 3804 AGGCAACAAACTGTTGGCCCCCGTGGCCCTGGCTGTGGGATCGATGGGAGCCCTCTTGT 3863

Qy 3769 GGGTGATTTCAACTACATTAGAAGGATCTTCCCCTCTGGAATGTCCACCAACATCCTAGA 3828

Db 3864 CGGTGACTTCAATTATATCCGGGCGCATCTTCCCCTTCTCGAAACGTGACCAGTATCTTGA 3923

Qy 3829 GCTGAGGAATAAAGATTTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCAC 3888

Db 3924 GTTACGAAATAAAGAGTTTAAACATAGCAACAGCCCGAGGACACAAAGTACTACTTGGCTGT 3983



Db 6141 GGGCACCAGGCGCAGGTGTTCTATAAGTACGGAATACTGTCCAAAGTTATCGGAGATCGT 6200  
QY 6106 CTATGACACCAAGAGTTCAGTTTCACTATGACGAGACGGCAGGCATGCTGAAGACCAT 6165  
Db 6201 CTACGACAGCACTGCCGTCACTTCGGCTATGACGAGACCACTTGGCGTCTTGAAGATGGT 6260  
QY 6166 CAACCTACAGATGAGGGCTTCACCTGACCACTCCGCTACCGTACAGATGGGCCCCCTGAT 6225  
Db 6261 GAATCTCCAAAGCGGGGCTTCTCCTGTACCATCAGGTACCGAAAGGTTCGGGCCCTCGT 6320  
QY 6226 TGACCGACAGATCTTCGGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTCCTGACTACA - 6284  
Db 6321 GGACAAGCAGATTACAGGTTCTCTGAGGAAGGCATGATCAACGCCAGGTTTCGATTACAC 6380  
QY 6285 --ACTATGACAAACAGCTTCGGGTGACCAAGCATGAGGCTGTGATCAACGAGACCCCACT 6342  
Db 6381 CTACCACGACACAGCTTCGGCATCGCCAGCATCAAGCCCGTCATCAGTGAGACTCCCT 6440  
QY 6343 GCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAGTT 6402  
Db 6441 TCCCGTTGACCTCTACCGCTACGATGAGATTCTGGCAAGGTGGAACACTTCGGCAGTT 6500  
QY 6403 TGGTGTCTATTTACTATGACATTAAACCAGATCATCAACCAAGCTGTATGACCCACCA 6462  
Db 6501 CGGGGTCTATCTACTACGACATCAACCAAGATCATCAACCACTGCGCTCATGACACTCAGCAA 6560  
QY 6463 GCATTTTGATGCTATGCGAGGATGAGGAAGTGCAGTATGAGATCTTCGGCTCGCTCAT 6522  
Db 6561 GCATTTTGACACCCATGGCGCATCAAGGAAGTGCAGTATGAGATGTTCCGGTCCCTCAT 6620  
QY 6523 GTACTGGATGACCGTCCAGTATGATTAACATGGGGCGAGTAGTGAAGAAAGGAGCTGAAGT 6582  
Db 6621 GTACTGGATGACCGTGCAATATGACAGTATGGCAGGGTCAATCAAGAGGGAACCTGAAACT 6680  
QY 6583 AGGACCTTACGCCAATACCACTCGCTACTCCTATGATGATGATGCTGACGGCCAGTGCA 6642  
Db 6681 GGGGCCCTATGCCAAACACCAAAAGTACACCTATGACTACGACGGGACGGCCAGTCCA 6740  
QY 6643 GACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCT 6702  
Db 6741 GAGTGTGGCGTCAATGACCGGCTACCTGGCGTTATAGCTATGACCTCAATGGGAACCT 6800  
QY 6703 GCATTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCA 6762  
Db 6801 GCACCTGCTAAACCCAGGAACAGTGTGCTCGCTCATGCGGTATGCTATGACCTCCGTGA 6860  
QY 6763 CCGCATCACTCGGCTGGGTGACGTGCAATACAAAGATGATGAGGATGGCTTCCTGAGGCA 6822  
Db 6861 CCGGATAACAGGCTAGGGACGTGAGTACAAATCGATGATGATGGCTATTTATGCCA 6920  
QY 6823 GCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCTGCTCATCAAGGCCTACACCG 6882  
Db 6921 GAGAGGATCTGACATCTTTGAATACAACTCCAAGGCGCTTCAACGAGAGCGTACAAACA 6980  
QY 6883 GGCTGGCAGCTGGAGTGTCAAGGTACCGTACCGTACGATGGCCTGGGCGCGGTGTCAGCAA 6942  
Db 6981 GGCCAGCGGTGGAGTGTGAGTACCGCTATGATGGCGTACCGCGCGGCTTCCTACAA 7040  
QY 6943 GAGCAGCCACAGCCACCACTCGAGTCTTCTATGACAGACCTGACCAACCCCAACAGGT 7002  
Db 7041 GACCAACCTGGGCCACCACTACAGTACTTCTATTCGACCTCCACCAACCCCAACAGTAT 7100  
QY 7003 CACCCACCTGTACAAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGAAGG 7062  
Db 7101 CACCCATGTTTACAACCACTCCAACCTGAGATCACTCTACTATGACCTCCAGGG 7160  
QY 7063 ACACCTCTTGCCATGGAGCTGAGAGTGGTGTGATGATTTTACATAGCTTGTGACAAAT 7122  
Db 7161 CCACCTCTTTGCCATGGAGAGCAGTAGTGGGAAGAGTACTATGTTGCTCAGATAACAC 7220  
QY 7123 CGGGACCCCTCTTGCTGTCTTTAGTGAACAGGTTTGTGATCAAGCAAAATCTGTACAC 7182

Db 7221 CGGGAATCCTCTGGCTGTTTTTTAGTATCAATGGCCTCATGATCAAGCAACTCCAATACAC 7280  
QY 7183 AGCCTATGGGAGATCTACATGGATACCAACCCCAACTTTTCAGATCATATAGGCTACCA 7242  
Db 7281 AGCCTATGGGAGATTTACTATGACTCCAATCCAGACTTTTCAGATGGTTCATCGGCTTCCA 7340  
QY 7243 TGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCCGGGAGATTATGATGT 7302  
Db 7341 CGGAGGCTCTACGACCCCTCACCAGCTCGTTCACTTTACGAGCGTGTATGATGAGT 7400  
QY 7303 GCTGGCCGAGCGTGGACTAGCCACAGACCGAGCTGTGGAAGCACCTTAGTAGCAGCAA 7362  
Db 7401 GCTGGCAGGACGGTGGACGTCCCCGACTACACCATGTGGAGAA--TGTGGGCAAGGA 7457  
QY 7363 CGTCATGCCCTTTTAATCTCTATATGTTCAAAAAACAACCCCATCAGCAACTCCCAGGA 7422  
Db 7458 GCCAGCCCTTCAACCTGTACATGTTCAAGAAACAACATCCACTCAGTAATGAGCTGGA 7517  
QY 7423 CATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACA 7482  
Db 7518 TTTAAAGAACTACGTGACAGACGTGAAGAGCTGGCTCGTATGTTTGGATTTTCAGCTCAG 7577  
QY 7483 CAACGTGATCCCTGGTTATCCCAACACAGACATGGATGCCATGGAAACCTCCTACGAGCT 7542  
Db 7578 CAACATCATCTCTGGATTCCCAAGAGCCAAAATGTATTTTGTGCTCCCCCTATGA--- 7634  
QY 7543 CATCCACACACAGATGAAAACGCAAGAGTGGGACAACAGCAAGTCTATCCTCGGGGTACA 7602  
Db 7635 -----ACTGTACAGAGAGCCAAAGTGAAGTGAAGTGGACAGCTCATACAGGTGTCCA 7685  
QY 7603 GTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGTCACTTTAGAACGGTTTGAACAGCTCTA 7662  
Db 7686 GCAGACAACAGAGAGGCATTAACCGGCCCTTTCTGGCTCTAGAAGGACAGGTCTATCTCTAA 7745  
QY 7663 TGGCTCCACAATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAAGTTTGCATCCAGGG 7722  
Db 7746 AAAGCTCCATGCAGGCATCCGAGAGAAAGC-----AGGCCACTGGTTTGTACGACCAC 7799  
QY 7723 CTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCACAGACAT 7782  
Db 7800 GCCCATCATCGGCAAGGCATCATGTTGCCATCAAAGAGGGCGGTGACCACAGGCGT 7859  
QY 7783 CATCAGTGTGGCCAAATGAGGATGGGCGAAGGTTGCTGCCATCTTGAACCATGCCCACTA 7842  
Db 7860 GTCTAGCATCGCCAGTGAGGACAGCCGCAAGGTAGCATCCGTGTTGAACAACGCTCTACTA 7919  
QY 7843 CCTAGAGAACTGCACCTTCAACCATGATGGGTGGATACCCATTACTTTGTGAAACCCAGG 7902  
Db 7920 CTTGGACAAGATGCACATACAGCATCGAGGGCAAGGACACACACTACTTCTGTAAGATCGG 7979  
QY 7903 ACCTTCAGAAGGTGACCTGGCCATCCTGGGCCCTCAGTGGGGGGCGGGAACCTTGGAGAA 7962  
Db 7980 TGCAGCGGACGGTGACCTGGTTACCTGGGGGACCAACATTTGGGCGCAAGGTGCTGGAGAG 8039  
QY 7963 TGGGTCACAGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTA 8022  
Db 8040 CGGGGTGAACGTGACCGTGTACAGCCCAACGCTGCTGTTGTAACGCGCAGGACTCGAAGTT 8099  
QY 8023 CACAGACATCCAGTCCAGTACGGGGCACTGTGCTTGAACACACAGTACGG----- 8073  
Db 8100 CACCAACATTTGAATTTCCAGTACTTCCAGCTGTGCTCAGCATACGTAACGGCCTCACCCC 8159  
QY 8074 GACAACTGTTGATGAGGAGAAGGACCGGTCCTGGAGCTGGCCCCGGCAGAGAGCCGTGCG 8133  
Db 8160 CGACACACTGGATGAAGAGAAGGCCGCTCTGGACCAAGCGCGACAGAGGGCCCTGGG 8219  
QY 8134 CCAAGCTGGGCCCGCGAGCAGAGAGACTGCGGGAAGGGAGGAAGCCCTGCGGGCTG 8193  
Db 8220 TACTGCTGGGCCAAGGAGCAGAGAAAGCCAGGGACCGGAGAGAGGGCAGCCGCTGTGTG 8279  
QY 8194 GACAGAGGGGGAAGCAGCAGTGTCTGAGCACAGGGCGGCTGCAAGGCTACAGCGGCTT 8253  
Db 8280 GACGGAGGGCGAGAAGCAGCACTCCTGAGCACGGGACCGGTGCAAGGTTTATGAGGGCTA 8339

QY 8254 TTTCTGATCTCTGTCGAGCAGTACCCAGAACTGTGACAGCGCCAAACAATCCACTT 8313  
Db 8340 TTACGTGCTTCCGGTGGAAACAGTACCCAGAGCTGGCAGACAGTAGCAGCAACATCCAGTT 8399  
QY 8314 CATGACACAGACGAGATGGCGCGGAGGTGACAGA 8348  
Db 8400 CTTAAGACAGAATGAGATGGGAAGAGGTAACAA 8434

RESULT 11  
US-09-808-602-7  
; Sequence 7, Application US/09808502  
; Patent No. US20020155115A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Heriman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same  
; FILE REFERENCE: 15966-697 CIP  
; CURRENT APPLICATION NUMBER: US/09/808,602  
; CURRENT FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/800,198  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 9826  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (280)..(8478)  
US-09-808-602-7

Query Match 40.5%; Score 3385; DB 9; Length 9826;  
Best Local Similarity 64.7%; Pred. No. 0;  
Matches 5393; Conservative 0; Mismatches 2790; Indels 154; Gaps 18;

QY 35 ATGACGTGAAGGAGAGGAGGACGCTTACCGCTGCTGACCGCGCGCGCGAGCGC 94  
Db 280 ATGATGTAAAGGACCGCGGACACCGCTCTTTGACCAGAGACGCTGTGGCAAGAGTGT 339  
QY 95 CGTACACACGAGCTCGTCCGGGACAGCGAGGAGGGGCAAG--CCCGCAGAAATCGTAC 151  
Db 340 CGTTACACAAGCTCCTCTCTGGACAGTGAGGACTGCCGGGTGCCACACAGAAATCCTAC 399  
QY 152 AGTCCAGCGAGACCCCTGAAGCCTACGACCGAGACGCCCGCCTAGCCTATGGCAGCCGC 211  
Db 400 AGTCCAGTGAGACTCTGAAGCCTATGACCATGACAGAGGATGCACATATGGAACCGA 459  
QY 212 GTCAAGGACATTGTGCCGAGGAGGCGGAGGAAATCTGCCGCACAGGTGCCAATTCACC 271  
Db 460 GTCAAGACCTCATCCACCGGAGTCAGATGAGTTTCTAGACAAGGAACCAACTTCACC 519  
QY 272 CTGCGGAGCTGGGGCTGGAAGAAGTAACGCCCTCACGGGACCCCTGTACCGGACAGAC 331  
Db 520 CTTGCCGAACCTGGGCATCTGTGA--GCCCTCCCCACACCGAAGCGGCTACTGTCCGAC 576  
QY 332 ATTGGCTGCCCCAATGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGAGGCT 391  
Db 577 ATGGGATCCTTCACCGAGGCTACTCCCTTAGCACAGGCTGTGACGCCGACTCCGACACC 636  
QY 392 GACACGGTGTCTCCCTGAGCACCCCGTGGCTCTGTGGGCGGAGCACACGCTCAGGG 451  
Db 637 GAGGAGGGATGTCTCCAGAACACGCCCATCAGACTGTGGGCGAGGGGATAAATCCAGG 696

QY 452 CGCAGTCTCTGCTGTCCAGCCGGGCGCAATTCCAATCTCACACTCACCGACACCGAGCAT 511  
Db 697 CGCAGTTCGGGCTGTCCAGTCTGTAATACTCGGCCCTTACCTGACTGACTCTGACAAC 756  
QY 512 GAAACACTGAGACTGATCATCCGGGCGGCC---TGCAGAACCAACGCGCGCTCCGGAC 567  
Db 757 GAAACAAATCAGATGATGAGAAGGTCGTCCCATTCACCTACATCTCGCCTAGTCTC 816  
QY 568 GCCGCGCGCGCTCTCGACGCCACACCCCAACCCAGACCAACGCGGCTCCATTAA 627  
Db 817 CTCCCATCTGCTCAGTGCCTAGCTCCATAATCTCCACAGTTAGTGCC-AGATGCC 875  
QY 628 CTCCCTGAACCGGGCAACTTCACGCCGAGGAGCAACCCAGACCCGCGCCACGGACCA 687  
Db 876 ATTGCTAGACAGCAACACCTCCCATCAAATCATGGACACCAACCTGATGAGGAATTC 935  
QY 688 CTCGCTCTCCGGAGAGCCCTCTCGCGGCGGCCAGGAGCTGCCACGCCAGAGAA 747  
Db 936 CCCCAATTCTACCTGCTCAGAGCATGCTCAGGGCCCCAGCAAGCCTCCAGCAGTGGCC 995  
QY 748 CTGGCTCTCAACAGCAACATCCCTTGAGACCAAGAACCTAGGCAAGCAGCATTCCT 807  
Db 996 TCCGAACACACAGCCAGTCGACTCTGAGGCC-----CCTCTCCC 1037  
QY 808 AGGACATTGCAGGACAACTCATTTGAGATGGACATTTCTGGCGCTCTCCGCCATGATG 867  
Db 1038 ACCCTCTCAACACACAGCTGTCCCATCACCCTCGTCCGCCAACTCCCTCAACAGGAA 1097  
QY 868 GGCTTACAGTGACGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCAC 927  
Db 1098 CTCACTGACCAATCGCGGAGTCAGATCCACGCCCGCGGCCAGGCCCAATGACCTGGC 1157  
QY 928 CACATCACAGGGTACCCACTGACGTCAGCTCCAGCACAGTGTACTCTCTCGCCCCGAC 987  
Db 1158 CACCACACAGAGTCCGTTACGCTTCAGGACAGCTGGGTGTAAACAGCAACGTGCCACT 1217  
QY 988 GCCCGGACAGCACTTCGCCCGCGCGCTTTAACTCAAGAACCCCTCAAGTACTGTAA 1047  
Db 1218 GGAGACCGGCACTTCCTCTTCAAGACCTCTCGGGGAGCACACCT-----T 1265  
QY 1048 CTGGAAGTGGCAGCCCTGAGCGCATCTGATCTCATCTCAGCCACTCTGGTCTATCTGTGC 1107  
Db 1266 GTTCAGCAGCTCTTCCCGGATACCTTTGACCTCAGGAACGGTTTACACGCCCCCGCC 1325  
QY 1108 ATACTTTGTGGCCATGCACCTGTTTGGCTTAACTGGCACCTGCAGCCGATGGAGGGCA 1167  
Db 1326 CCGCTGTGCCAGGAATACTTTCTCCA-----GGAAGGCTTTCAA 1367  
QY 1168 GATGTATGAGATCAGGAGGACACAGCCAGCAGTGGCTGTGCCAACCGACGTCTCC-C 1226  
Db 1368 GCTGAAGAAGCCCTCCAAATACTGCAGCTGGAATGTGTGCTGCCCTCTCGCCATTTGCCG 1427  
QY 1227 TATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGGAACACAG 1286  
Db 1428 GGCCCTCTCTTGGCTATTTTGTGGCTATTTTCATAGTGCCTGGTCTGTGAAAAACAG 1487  
QY 1287 AAGGAAGCCCGAGTAGTTTCTTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATG 1346  
Db 1488 CAG-----CATAGACAGTGGTGAAGCAGAAG 1513  
QY 1347 TGGGAAGCGGAGCCCTCCAGAAAGATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTC 1406  
Db 1514 TTGGTCGGCGGGTAACACAAGAGTCCACAGGGGTGTTTGGAGGTCACAAATTCACA 1573  
QY 1407 TAGACCATCTGTGCACTGAAATCAATGTGTCTCTGGGAAGGCGAGCCCTGGTTGGCA 1466  
Db 1574 TCAGTCAGCCCCAGTTCTTAAAGTTCAACATCTCCCTCGGGAAGGACGCTCTCTTTGGTG 1633  
QY 1467 TTTATGGCAGAAAAAGCCCTCCCTCTTACATACACAGTTTCACTTTGTGGAGCTGCTGG 1526  
Db 1634 TTTACATAAGAAGAGGACTTCCACCATCTCATGCCAGTATGACTTTCATGGAACGTCTGG 1693



Db 3833 CCAGCTGCAACGGCCTTGCTGAAGGCAACAAGCTGCTGGCCCCAGTGGCTCTGGCTGTTG 3892

Qy 3747 GCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATATTAGAAGGATCTTCCCCTCTG 3806

Db 3893 GAATCGATGGGAGCCTCTATGTGGGTGACTTCAATTACATCCGACGCATCTTCCCTCTC 3952

Qy 3807 GAAATGTCACCAACATCCTTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCACAGTCCAG 3866

Db 3953 GAAATGTGACCAGCATCTTTGAGATTACGAAATAAAGATTTAAACATAGCAACAACCCAG 4012

Qy 3867 CACACAAATACTACCTGGCCACAGACCCCCATGATGGGGCCGTCTTCCCTTTCTGACAGCA 3926

Db 4013 CACACAAGTACTACTTTGGCAGTGGAACCCCGTGTCCGGCTCGCTCTACGTGCCACACCA 4072

Qy 3927 ACAGCCGGCGGGTCTTTAAATAAAGTCCACTGTGTGGTGAAGGACCTTGTCAAGAACT 3986

Db 4073 ACAGCAGGAGAACTACCGCGTCAAGTCTCTGAGTGGAAACCAAGACCTGGCTGGGAATT 4132

Qy 3987 CTGAGGTGGTTGCGGGACAGGTGACAGTGCCCTCCCTTTGATGACACTCGGTGCGGGG 4046

Db 4133 CGGAAGTTGTGGCAGGGACGGGAGAGCAGTGTCTACCCCTTGATGAAGCCCGTGGGGG 4192

Qy 4047 ATGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCAGGGGTATTACAGTGGACAAGT 4106

Db 4193 ATGAGGGGAAGGCCATAGATGCAACCCCTGATGAGCCGAGAGGTATTGCAGTAGACAAGA 4252

Qy 4107 TTGGGCTGATCTACTTTCGTGGATGGCACCATGATCAGACGCATCGATCAGAAATGGGATCA 4166

Db 4253 ATGGGCTCATGTACTTTGTGATGCCACCATGATCCGGAAGTTGACCAGAATGGAATCA 4312

Qy 4167 TCTCAACCTTGCTCGGCTCTAATGATCTCACATCAGCCCGGCCACTCAGCTGATTTCTG 4226

Db 4313 TCTCAACCTTGCTGGCTCCCAATGACCTCACTGCCGTCCGGCCGCTGAGCTGTGATTCCA 4372

Qy 4227 TCATGGATATTTCCAGGTAAGACTGGAGTGGCCCCACAGACTTAGCCATCAACCCAAATGG 4286

Db 4373 GCATGGATGTAGCCCAAGTTTCGTCTGGAGTGGCCACAGACCTTGTGTCAATCCCATGG 4432

Qy 4287 ACAACTCACTTTATGTCCCTGACAAACAATGTGGTCTTCGAAATCTCTGAAAAACCCACAGG 4346

Db 4433 ATAACCTCTTGTATGTTCTAGAGAAACAATGTCACTCCTTCGAATCACCGAGAACCAACG 4492

Qy 4347 TGCGCATTTGTCGCCGGGAGGCCCATGCCATGCCAGTCCCTGGCATTTGACCACTTCCCTGC 4406

Db 4493 TCAGCATCATTTGCGGACGCCCCCATGCATGCCAAGTTCCTGGCATTGACTACT--CAC 4549

Qy 4407 TAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCACCGCTTTGGCTGTTTCACACA 4466

Db 4550 TCAGCAAACTAGCCATTCACTCTGCCCTGGAGTCAGCCAGTGCCATTGCCAATTTCTCACA 4609

Qy 4467 ATGGGCTCCTGTATATTGCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTCA 4526

Db 4610 CTGGGTCCTCTACATCACTAGACACAGATGAGAAGAAGATTAACCGTCTACGCCAGGTAA 4669

Qy 4527 CCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCCCCAGTGGCTGTGACTGTAAAAAATG 4586

Db 4670 CAACCAACGGGGAGATCTGCCTTTTAGCTGGGGCAGCCTCGGACTCGGACTGCAAAAACG 4729

Qy 4587 ATGCCAACTGTGATTGTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATA 4646

Db 4730 ATGTCAATTGCAACTGCTATTTCAGGAGATGATGCCCTACCGACTGATGCCATCTTGAATT 4789

Qy 4647 CCCCATCTTCTTGGCTGTGTGCTGATGGGGAGCTTACGTGGCCGACCTTGGGAACA 4706

Db 4790 CCCCATCATCTTAGCTGTAGTCTCCAGATGGTACCATTACATTGCAGACCTTGGAAATA 4849

Qy 4707 TCCGAATTCGGTTTATCCGGRAGAACAAAGCCTTTCCTCAACACCCAGAACATGTATGAGC 4766

Db 4850 TTCGATCAGGGCGGTTCAGCAAGAACAAAGCCTGTTCTTAATGCCTTCAACCAAGTATGAGG 4909

Qy 4767 TGTCTCACCAATTGACCAGGAGCTCTATCTGTTTTGTATACCACCGGCAAGCACTGTACA 4826

Db 4910 CTGCATCCCCCGGAGAGCAGGAGTTATATGTTTTCAACGCTGATGGCATCCACCAATACA 4969

Qy 4827 CCCAAAGCCTGCCACACAGGAGACTACCTGTACAACCTTCACTACATCTGGGACGGCGACA 4886

Db 4970 CTGTAGCCTGGTGACAGGGGAGTACTTGTACAATTTACATATAGTACTGACAATGATG 5029

Qy 4887 TCACACTCATCACAGACAACAATGGCAACATGGTAAATGTCCGCCGAGACTCTACTGGGA 4946

Db 5030 TCACTGAATTGATTGACAATAATGGGAATTCCTCTGAAGATCCGTGGGACAGCAGTGGCA 5089

Qy 4947 TGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACA 5006

Db 5090 TGCCCCCTCACCTGCTCATGSCCTGACAACACAGATCATCACCTCACCGTGGCACCCTATG 5149

Qy 5007 GTGCACTCAAGAGTGTGACCCACACAAGGACACGAGTTGGCCATGATGACATACCATGGCA 5066

Db 5150 GAGGCTTCAAAAGTCGTGTCCACACACAGAACCTGGAGCTTGGTCTCATGACCTATGATGGCA 5209

Qy 5067 ATTCGGGCTTCTGGCAACCCAAAAGCAATGAAAAACGATGGACACAACATTTTATGAGTACG 5126

Db 5210 ACATGGGCTCCTGGCCACCAAGAGCGATGAAACAGGATGGACGACTTTCTATGACTATG 5269

Qy 5127 ACAGCTTTGGCCGCTTGACAATAATGTGACCTTCCCTACTCTGCCAGGTGAGCAGTTTCCGAA 5186

Db 5270 ACCACGAAGCCGCTTGACCAACGTGACGGCCCCACGGGGTGGTAAACAGTCTGCACC 5329

Qy 5187 GTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTCCA--GCAAGGATGATGTCA 5243

Db 5330 GGGAAATGGAGAAATCTATTACCATTGACATTGAGAACTTCAACCGTGATGATGACGTCA 5389

Qy 5244 CCATAACCAACCAACCTGTCTGCCTCAGGCGCCTTCTACACACTGCTGCAAGACCAAGTCC 5303

Db 5390 CTGTCAATCAACCAACCTCTCTTCAGTAGAGGCCCTCCTACACAGTGGTACAAGATCAAGTTC 5449

Qy 5304 GGAACAGCTACTACATCGGGGCCGATGGCTCCTTTCGGGCTGCTGCTGGCCAAACGGCATGG 5363

Db 5450 GGAACAGCTACAGCTCTGTATAATAATGGTACCCTGAGGGTCAATGCTAATGGGATGG 5509

Qy 5364 AGTGGCGCTGCAGACTGAGCCCACTTGTGGTGGCACCGTCAACCTGGTGGAGTGGCGCAGCGCA 5483

Db 5570 GCTGCAACATCTCCCTGCCTATGGAGAATGGCTTAACTCCATTGAGTGGCGCCTAAGAA 5629

Qy 5484 AAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTTGGCGCCGCTGCGGTCGCAACAACCGAA 5543

Db 5630 AGGAACAGATTAAAGGCAAAAGTCACCATCTTTGGCAGGAAGTCCGGGTCCAATGGAAGAA 5689

Qy 5544 ATCTCCTATCTCTGGACTTTTGATCGCGTAAACACGCAAGAGAAGATCTATGATGACCACC 5603

Db 5690 ATCTCTTGTCATTGACTATGATCGAAATATTTCGGACTGAAAAGATCTATGATGACCACC 5749

Qy 5604 GCAAGTTCACCCCTTCGGATTCTGTACGACCAAGCGGGGGCGGCCAGCCCTCTGGTCACCCA 5663

Db 5750 GGAAGTTCACCCCTGAGGATCATTTATGACCAGGTGGCGCCGCCCTTCTCTGGCTGCCCA 5809

Qy 5664 GCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTTGGGGGTTACATTGCTGGCATCC 5723

Db 5810 GCAGCGGCTGGCAGCTGTCAACGTGTCACTACTTTCTTCAATGGGCGCCTGGCTGGGCTTC 5869

Qy 5724 AGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCAGGGCGCGCATCACATCCAGGA 5783

Db 5870 AGCGTGGGGCCATGAGCGAGAGGACAGACATCGACAAGACGCCGCATCGTGTCCCGCA 5929

Qy 5784 TCTTCGCTGATGGGAAGACATGGAGCTACACATACTTAGAGAAGTCCATGGTGTGCTAC 5843

Db 5930 TGTTCGCTGACGGGAAAGTGTGAGCTACTCCTACTCTTGACAAGTCCATGGTCTCCTGTC 5989

Qy 5844 TACACAGCCAGAGGCAGTATATCTTTGAGTTTCGACAAGAATGACCGCCTCTCTTCTGTGA 5903

Db 5990 TTCAGAGCCCAACGTCAGTATATATTGAGTATGACTCCTCTGACCGCCTCCTTGCCGTCA 6049



Db 8189 GCATCCGCTATGGCCTCACCCCCCGACACCCTGGACGAAGAGAGGCCCGCTCCTGGACC 8248  
Qy 8112 TGGCCCGGCAGAGAGCCGTGCGCCCAAGCGTGGGCCCCGCGAGCAGCAGAGACTGCGGGGAA 8171  
Db 8249 AGCGGAGACAGAGGGCCCTGGGCACGGCCTGGGCCCAAGAGCAGCAGAAAGCCAGGGACG 8308  
Qy 8172 GGGAGGAAGGCCTGCGGGCTGGACAGAGGGGGAGAGCAGAGGTGCTGAGCACAGGGC 8231  
Db 8309 GGAGAGAGGGGAGCGCCCTGTGGAAGAGGGCGAGAGAGCAGACTTCTGAGCACCGGGC 8368  
Qy 8232 GGGTGAAGGCTACGACGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTGAC 8291  
Db 8369 GCGTGAAGGGTACGAGGGATATTACGTGCTTCCCGTGGAGCAATACCCAGAGCTTGAC 8428  
Qy 8292 ACAGCGCCCAACAATCCACTTCATGAGACAGAGCGGCGGAGGTGACAGA 8348  
Db 8429 ACAGTAGCAGCAACATCCAGTTTAAAGACAGATGAGATGGGAAGAGGTAAACAAA 8485

RESULT 12  
US-09-800-198-7  
; Sequence 7, Application US/09800198  
; Publication No. US20030087816A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Cornie AM  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herrmann, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishna  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Rastelli, Luca  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-697  
; CURRENT APPLICATION NUMBER: US/09/800,198  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 9826  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (280)..(8478)  
US-09-800-198-7

Query Match 40.5%; Score 3385; DB 10; Length 9826;  
Best Local Similarity 64.7%; Pred. No. 0;  
Matches 5393; Conservative 0; Mismatches 2790; Indels 154; Gaps 18;  
Qy 35 ATGACGTGAAGGAGAGGAGCCCTTACCGCTCGCTGACCCCGCGCGGACCGCGAGCGC 94  
Db 280 ATGATGTAAGGACCGGCGACACCGCTCTTTGACCAGAGGACGCTGTGGCAAGAGTGT 339  
Qy 95 CGTACACCAAGCTCGTCCGCGGACAGCGAGGAGGGGCAAG---CCCGCAGAAATCGTAC 151  
Db 340 CGTACACAAGCTCCTCTCTGGAAGTGGAGGACTGCGGGTGCCACACAGAAATCCTAC 399  
Qy 152 AGTCCAGCGAGACCCCTGAAGGCTTACGACAGGACCGCCCGCTAGCCCTATGCGCAGCCGC 211  
Db 400 AGTCCAGTGAGACTCTGAAGGCTTATGACCATGACAGCAGGATGCATATGGAACCGA 459  
Qy 212 GTCAAGGACATTGTGCCGAGGAGCCGAGGAATTCTGCCGACAGGTGCCAATTCACC 271  
Db 460 GTACAGACCTCATCCACCGGAGTACAGTGTGCTTCCCTAGACAAGGAACCAACTTCACC 519  
Qy 272 CTGCGGAGCTGGGCTGGAAGAAGTAACGCCCCCTCAGGGACCCCTGTACCGGACAGAC 331  
Db 520 CTTGCCGAACCTGGGCTCTGTGA---GCCCTCCCCACACCGAAGCGGCTACTGCTCCGAC 576

Qy 332 ATTGGCCTGCCCCAATGCGGCTACTCCATGGGGGCTGGCTTCTGATGCCGACATGGAGGCT 391  
Db 577 ATGGGATCCTTTCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCCGACTCCGACACC 636  
Qy 392 GACACGGTGTGTCCCTGAGACACCCCGTGTGTGTGGGCGGAGCACACGGTCAAGG 451  
Db 637 GAGGAGGGATGTCTCCAGAACACGCCATCAGACTGTGGGCGAGAGGATAAAATCCAGG 696  
Qy 452 CGAGTCTCTGCTGTCCAGCGCGGCCAATTCCAAATCTCACACTCACCGACACCGAGCAT 511  
Db 697 CGCAGTTCGGCCTGTCCAGTGTGTAATACTCGGCCCTTACCCTGACTGACTCTGACAA 756  
Qy 512 GAAACACTGAGACTGATCATCCGGGCGGCC---TGAGAAACACACGCGCGGCTCCGGAC 567  
Db 757 GAAACAAATCAGATGATGAGAACGGTGTGTCCCATTCACCTACATCTCGCCTAGTCTC 816  
Qy 568 GCCGCGCGCGCTCTCTCGACGCGCCACACCCCAACACGACACCGCGGCTCCATTA 627  
Db 817 CTCCCATCTGTCTCAGTGTGCTAGCTCCATAATCTCTCCACGATTAGTGTCC-AGATGCC 875  
Qy 628 CTCCTGAACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCCCCACCGGACCA 687  
Db 876 ATTGTAGACAGCAACACCTCCCATCAAATCATGGACACCAACCTGTATGAGGAATTCTC 935  
Qy 688 CTCGCTCTCCGGAGAGCCCTGCGGGCGGCGCCAGGAGCTGCCACGCGCCAGGAGAA 747  
Db 936 CCCCAATTACATCTGTCTCAGAGCATGTCTAGGGCCCCAGAACGCTCCAGCAGTGGCCC 995  
Qy 748 CTGGCTGTCAACAGCAACATCCCTCTGGAGACAGGAACTTAGGCAAGCAGCATTCTCT 807  
Db 996 TCCGAACCAACACAGCCAGTCTGAGGCC-----CCTCTCCC 1037  
Qy 808 AGGGACATTGCAGGACAACTTCATTGAGATGGACATTCTCGGCGCTCCCGCCATGATGG 867  
Db 1038 ACCCCCTCAACACACAGCTGTCTCCATCACCACCTCGTCGCCAACTCCCTCAACAGGAA 1097  
Qy 868 GGCTTACAGTACGGGCACCTTCTTCAAGCCTGGAGGACCTTCCCGCTCTTCTGTGAC 927  
Db 1098 CTCACTGACCAATCGGGGAGTCTGATCCACGCGCGCGCCAGCGCCCAATGACCTGGC 1157  
Qy 928 CACATCACCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCT 987  
Db 1158 CACCACACAGAGTCCGTTAGCTTCAGGACAGCTGGGTGTAAACAGCAACGTGCCACT 1217  
Qy 988 GCCCGCAGCACCTTCCGCGCGGCGCTTTAACTCAAGAACGCGCTTCAAGTACTGTAA 1047  
Db 1218 GGAGACCGGCACTTCTCTTCAAGACCTCTCGGGGAGCACACCT-----T 1265  
Qy 1048 CTGGAAGTGCAGCCCTGAGCGCCATCGTCACTCTCAGCACTCTGGTCACTCTGTGGC 1107  
Db 1266 GTTACAGCAGCTTTCGCCGGGATACCTTTGACCTCAGAACGCTTACACGCCCCCGCC 1325  
Qy 1108 ATACTTTGTGGCCATGCACCTGTGTGGCTTAACTGGCACTGCAGCCGATGGAGGGCA 1167  
Db 1326 CCGCTGTCTGCCAGGAATACTTTCTCCA-----GGAAGGCTTTCAA 1367  
Qy 1168 GATGTATGAGATCACGGAGGACACAGCCAGCAGTGTGGCTGTGCCAACCGACTCTCC-C 1226  
Db 1368 GCTGAAGAAGCCCTCCAAATACTGACGTGGAAATGTGTGCTGCTCTCCGCCATTGCCGC 1427  
Qy 1227 TATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGCAAGAACCCACAG 1286  
Db 1428 GGCCCTCTCTTGGCTATTTTGTGGCGTATTTTCATAGTGTGCTGTGTTGAAAAACAG 1487  
Qy 1287 AAGGAAGCCCGAGTAGTTTCTTCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATG 1346  
Db 1488 CAG-----CATAGACAGTGTGAAGCAGAG 1513  
Qy 1347 TGGGAAGCGGAGCCCTCCAGAAAGATTCTCTCTGGCATTCTTCTGGAGATCTCAAGTGTCA 1406  
Db 1514 TTGGTCGGGGGTAAACAAAGAGTCCCAAGGGGTGTTTGGAGGTCACAAATTCACA 1573  
Qy 1407 TAGACCATCCTGTGCATCTGAAATTCAATGTGTCTCTGGGAAAGGACGCCCTGGTTGGCA 1466

Db 1574 TCAGTCAGCCCCAGTCTCTTAAAGTTCAACATCTCCCTCGGGAAGGACGCTCTCTTTGGTG 1633  
QY 1467 TTTATGGCAGAAAAGCCCTCCCTCCTTACATACACAGTTTGTACTTTGTGGAGCTGCTGG 1526  
Db 1634 TTTACATAAGAAGAGGACTTCCACCATCTCATGCCAGTCCAGTATGACTTCATGGAACGTCGG 1693  
QY 1527 ATGGCAGGAGGCTCCTAAACCAGGAGGCGGAGCCTAGAGGGACCCCGCGCCAGTCTC 1586  
Db 1694 ACGGA-----AGGAGAAAGTGGAGTGTGTTGAGTCTCCAGGGAACGCC 1738  
QY 1587 GGGGAACGTGCCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAA 1646  
Db 1739 GGAGCATACAGACCTTGGTTCAGAAATGAAGCCGTGTTTGTGCAGTACCTGGATGTGGCC 1798  
QY 1647 TCTGGCACTTGGCTTTTACAATGACGGAAGGAGTCAAGAGTGTCTTCTTCACCA 1706  
Db 1799 TGTGGCATCTGGCCTTCTACAATGATGGAAGAACAAAGAGATGGTTTCTTCAATACTG 1858  
QY 1707 CTGCCATTGAGTCGTGGATAACTGCCCCAGCAACTGCTATGGAATGGTGACTGCATCT 1766  
Db 1859 TTGTCTTAGATTGAGTCAGGACTGTCCACGTAACTGCCATGGGAATGGTGAATGTGTG 1918  
QY 1767 CTGGGACCTGCCACTGCTTCTCTGGGTTTCTTGGCCCCGACGTGGCAGAGCCTCCTGCC 1826  
Db 1919 CCGGGGTGTCTACTGTCTCCAGGATTTCTAGGAGCAGACTGTGCTAAAGCTGCCTGCC 1978  
QY 1827 CCGTGTCTGTAGCGGAATAGGCCAATACATGAAGGCAGATGCTGTGTCACACAGTGGCT 1886  
Db 1979 CTGTCTGTGAGTGGGAATGGACAATATCTAAAGGACGTGCCAGTGTCTACAGCGGCT 2038  
QY 1887 GGAAGGGCTGAGTCGATGTGCCCAATACATGAAGGCAGATGCTGTGTCACACAGTGGCT 1946  
Db 2039 GGAAGGTGCAGATGCGACGTGCCCATGAATCAGTGCAATCGATCGATCCTTCTCGGGGGCC 2098  
QY 1947 ATGGCACCTGTACAGGGCACCTGTGATCTGCAACCCCTGGCTACAAAGGCGAGAGCTGTG 2006  
Db 2099 ACGGCTCCTGCATTGATGGGAACGTGTGTCTGTCTGTGGCTACAAAGGCGAGCACGTGTG 2158  
QY 2007 AGGAAGTGGATGACAGCCCACTTCTCCCGGACACCGGGCTTGCAGCTGTGACCCAA 2066  
Db 2159 AGGAAGTTGATTGCTTGGATCCCACCTGTCCAGCCAAGGAGTGTGTGAATGGAGAA 2218  
QY 2067 GCCATTGCTTTGTGGATGGGAGGCACCAACTGCGAGACCCCGAGGCTTGCAGCTGTGACCCAA 2126  
Db 2219 GCCTGTGACGCCCTGGCTGGGGTGGTCTGAACTGTGAGCTGGCAGGGTCCAGTGCCTCAG 2278  
QY 2127 ACCAGTGTTCAGGCAACGGAACCTTCTCCCGGACACCGGGCTTGCAGCTGTGACCCAA 2186  
Db 2279 ACCAGTGCAGTGGGCATGGCACGTACCTGCTGACACGGGCTTGCAGCTGCGATCCCA 2338  
QY 2187 GCTGGACTGGACAGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGCG 2246  
Db 2339 ACTGGATGGGTCCGACTGCTCTGTTGAAGTGTGCTCAGTAGACTGTGGCACTCACGGCG 2398  
QY 2247 TGTGCGTAGGGGSCACCTGCCGCTGCGAGGATGGCTGGATGGGGCAGCCCTGCGACCCAGC 2306  
Db 2399 TCTGCATCGGGGAGCCTGCCGCTGTGAAGAGGGCTGGACAGGGCAGCGTGTGACCCAGC 2458  
QY 2307 GGGCCTGCCACCCCGCTGTGCCGAGCATGGGACCTGCCCGCAGCGCAAGTGGAGTGCA 2366  
Db 2459 GCGTGTGCCACCCCGCTGCATTGAGCAGGGACCTGTAAAGATGGCAATGTGAATGCC 2518  
QY 2367 GCCCTGGCTGGAATGGCGAACACTGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAG 2426  
Db 2519 GAGAGGGCTGGAATGGTGAACACTGCACCATTTGGTAGGCAACCGCAGGCCACCGAAACAG 2578  
QY 2427 AGGGTTGCCCTGGTTGTGCAATGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGCC 2486  
Db 2579 ATGGCTGCCCTGACTTGTGCAACGGTAAACGGGAGATGCACACTGGGTGAGAACAGCTGGC 2638  
QY 2487 ACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCCT 2546

Db 2639 AGTGTGTCTGCCAGACCGGCTGGAGAGGCCCCGGATGCAACGTTGCCATGGAACTTCCT 2698  
QY 2547 GCGGTGACAGCAAAAGACAATGATGGAGATGGCCCTGGTGGACTGCATGGACCCCTGACTGCT 2606  
Db 2699 GTGCTGATAACAAGGATAATGAGGGAGATGGCCCTGGTGGATTGTTGGACCCCTGACTGCT 2758  
QY 2607 GCCTCCAGCCCTGTGCCATATCAACCCGCTGTGCCCTTGGCTCCCTAACCCCTCTGGACA 2666  
Db 2759 GCCTGCAGTCAGCCTGTCAAGAACAGCCTGCTTGCCTGGGGGTCCCGGACCCACTGGACA 2818  
QY 2667 TCATCCAGGAGACACAGGTCCTGTGTGCACAGCAGAACCTACACTCCTTCTATGACCGCA 2726  
Db 2819 TCATTGAG-----CAGGGCCAGACGGATTGGCCCGCAGTGAAGTCTTCTATGACCCGTA 2872  
QY 2727 TCAAGTTCTCGTGGCAGGACAGCAGCAGACATAATCCCCGGGAGAACCCCTTTTGATG 2786  
Db 2873 TCAAGCTCTTGGCAGGCAAGGATAGCACCCACATCATTCCTGGAGAGAACCCCTTCAACA 2932  
QY 2787 GAGGGCATGCTGTGTATTTCGTGGCCAAAGTATGACATCAGATGGAACCCCTGGTTG 2846  
Db 2933 GCAGCTTGGTTTCTCTCATCCGAGGCCAAAGTAGTAACACAGATGGAACCTCCCTGGTCG 2992  
QY 2847 GTGTGAACATCAGTTTGTCAATAAACCTCTCTTTTGGATATACATCAGCAGGCAAGATG 2906  
Db 2993 GTGTGAACGTGTCTTTGTCAAGTACCCAAAATAACGGCTACACATCACCCGCCAGGATG 3052  
QY 2907 GCAGCTTTGACTTGGTGAACAAATGGCGGCATCTCCATCATCTCGCGGTTTCGAGCGGCAC 2966  
Db 3053 GCAGCTTCGACCTGATCGCAATGGAGTGCTTCTTGACTCTACACTTTGAGCGAGCCC 3112  
QY 2967 CTTTCATCACACAGGAGCACACCCCTGTGGCTGCCATGGGATCGGTTCTTTGTCAATGAAA 3026  
Db 3113 CGTTTCATGAGCCAGGAGCGCACCTGTGTGGCTGCCGTGGAACAGCTTTACGCCATGGACA 3172  
QY 3027 CCATCATCATGAGACATGAGGAGAAATGAGATTTCCAGCTGTGACCTGAGCAATTTTGCCC 3086  
Db 3173 CCCTGGTATGAAGACCGAGGAGAACTCCATCCCGAGCTGTGACCTCAGTGGCTTTGTGCC 3232  
QY 3087 GCCCAACCCAGTCGTCTCTCCATCCCGACTGACGTCTTCGCGAGCTCCTGTGCAGAGA 3146  
Db 3233 GGCCTGATCCCAATCATCATCTCCTCCCGACTGTCCACCTTCTTAGTGCTGCCCTGGGC 3292  
QY 3147 AAGGCCCCATTGTCCCGAAATTCAGGCTTTGCAGGAGGAAATCTCTATCTGTGGCTGCA 3206  
Db 3293 AGAATCCCATCGTGCCTGAGACCCAGGTTCTTCATGAAGAAATCGAGCTCCCTGGTTCCA 3352  
QY 3207 AGATGAGGCTGAGCTACCTGAGCAGCCGACCCCTGGCTACAAATCTGTCTCTGAGGATCA 3266  
Db 3353 ATGTGAAACCTTCGCTATCTGAGCTCTAGAAGCTGACGGGTACAAGTCACTGCTGAAGATCA 3412  
QY 3267 GCCTCACCCACCGACCATCCCTTCAACCTCATGAAGTGCACCTCATGGTAGCGGTGG 3326  
Db 3413 CCATGACCCAGTCCACAGTGCCCTGAAACCTCATTAGGGTTCACTGATGGTGGCTGTCG 3472  
QY 3327 AGGGCCGCTCTTTCAGGAAGTGGTTTCGCTGCAGCCCCCAGACCTGTCTCTATTATTTCATTT 3386  
Db 3473 AGGGGCATCTCTTCCAGAAAGTCAATTCAGGCTTCTCCCAACCTTGGCTCCACCTTCATCT 3532  
QY 3387 GGGACAAGACAGACGTCTACAAACAGAGGTGTTTCGCTGCAGCCCCCAGACCTGTCTCTATTTTCCG 3446  
Db 3533 GGGACAAGACAGATGCGTATGGCCAAAAGGGTGTATGGACTCTCAGATGCTGTGTGTCTG 3592  
QY 3447 TGGGTTATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTGC 3506  
Db 3593 TCGGGTTTGAATATGAGACCTGTCCAGTCTAATTTCTCTGGGAGAAAAGACAGCCCTCC 3652  
QY 3507 TGCAGGGCTATGAAATGACGCGTCCAAGCTTGGAGGATGGAGCTAGAGCTAGACAAACATCATG 3566  
Db 3653 TTCAGGGATTGAGCTGGACCCCTCCAACTCGGTGGTGGTCTCTAGACAAACACACACA 3712  
QY 3567 CCCTCAACATTCAAAAGTGGTATCTTCGACAAAAGGGAATGGGAGAACCCAGTTTGTGTCTC 3626  
Db 3713 TCCTCAATGTATAAAGTGAATCTCTACACAAAGGCACCTGGGAAAAAACAGTTTCCTGACCC 3772







QY 1250 TTAGAGACCCCTGACAGGAAAGGCAAGGAACCAAGAAAGGCCAGTAGTTTCTTT 1309  
Db 1424 CCCTGGTCATTGAAAA-----1439  
QY 1310 CCAGAGGACAGTTTCATAGATTCTGGAGAAATTGATGTGGAAAGGCGAGCCTCCCAGAAG 1369  
Db 1440 -----ACAGCAGCATAGACAGTGGCGAAGCAGAGAAGTTGGTCGGCGGGTGACACAGGAA 1492  
QY 1370 ATTCTCCTGGCACTTCTGGAGATCTCAAGTGTTCATAGACCATTCTGTGCACTCTGAAA 1429  
Db 1493 GTCCACACAGGGGTGTTTGGAGGTCCCAGATTCAACATCAGTCAGCCTCAATTCTTTAAAG 1552  
QY 1430 TTCAATGTGTCTCGGAAAGGACGCCCTGGTGGCAATTATGGCAGAAAAAGCCCTCCCT 1489  
Db 1553 TTCAACATCTCCCTGGCAAGGATGCCCTCTTCGGTGTCTATATAAGGAGAGACTACCA 1612  
QY 1490 CCTTCACATACACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAG 1549  
Db 1613 CCGTCTCATGCCAGTATGACTTCATGGAACGCCCTGGATGGAA-----AG 1657  
QY 1550 GAGCGCGGAGCCTAGAGGGGACCCCGCGCAGTCTCGGGGAATGTGCCCCCTCCAGC 1609  
Db 1658 GAGAAATGGAGCGTGGTCGAGTCGCCCGCAGGAAACGCCGAGCATCCAGACTCTGGTGCAG 1717  
QY 1610 CATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTACAAAT 1669  
Db 1718 AACGAGGCTGTGTGTGCAGTACTTGGATGTGGCCTGTGGCACTGGCCTTCTACAAAT 1777  
QY 1670 GACGGAAAGGAGTCAGAAAGTGGTTTCTTCTCACCACTGCCAATTGATCGGTGGATAAC 1729  
Db 1778 GACGGCAAGGACAAGGAGATGGTCTCCTTCAACACTGTTGTCTTAGATTTCAGTGCAGGAC 1837  
QY 1730 TGCCCCAGCACTGCTATGGCAATGTGACTGCATCTCTGGGACTGCCACTGCTTCTCTG 1789  
Db 1838 TGTCCACGGAACTGTACGGGAACGGTGAATGCGTGTCTGACTGTGTCACTGTTTCCCA 1897  
QY 1790 GGTTCCTGGGCCCGACTGTGGCAGAGCCTCTGCCCGTGTCTGTAGCGGAAATGGC 1849  
Db 1898 GGATTCTTAGTCAGACTGTGCTAAAGCTGCCCTGCTGCTGTCAGCGGAAATGGA 1957  
QY 1850 CAATACATGAAGGCAGATGCTTGTGCCACAGTGGCTGGAAAGCGCTGAGTGCAGTGTG 1909  
Db 1958 CAGTATTCTAAAGGAACGTGCCAGTGTACAGCGGCTGGAAGGTGCAGAGTGTGATGTG 2017  
QY 1910 CCCACCAACCACTGTATCGATGTGGCTGCGAGCAACCATGGCACTGCATCACGGGCACC 1969  
Db 2018 CCTATGAACCAATGTATCGATCCTTCTGTGGGGCCATGGCTCCTGCTGATGGGAAC 2077  
QY 1970 TGCATCTGCAACCCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCC 2029  
Db 2078 TGCGTGTGTGCTGGCTACAAGGGCGAGCACTGTGAGGAAGTGTGCTTGGATCCT 2137  
QY 2030 ACATGTTTCAGSCGGGGTGTCTGCGTGAGAGGCGAATGCCATTGCTTTGTGGGATGGGA 2089  
Db 2138 ACCTGCTCCAGCCTAGTGTCTGTGTGAATGGAGAGTGTCTATGACGCCCCGGCTGGGT 2197  
QY 2090 GGCACCAACTGAGACCCCGGAGCATGCTTAGACCAGTGTTCAGGCCACGGAACC 2149  
Db 2198 GGTCTCAACTGTGAGCTGGCGAGGTCCAGTGCACGACCAAGTGTAGTGGGCATGGCACT 2257  
QY 2150 TTCTCTCCGACACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTCT 2209  
Db 2258 TACCTCCCTGACTCCGGCTCTGCACTGTGATCCGAACGTGGATGGGTCCCGACTGCTCT 2317  
QY 2210 ATCGAGATCTGTGCTGCCGACTGTGTGGCCCATGGCGTGTGGTAGGGGGCACTTCCCGC 2269  
Db 2318 GTTG--TGCTCAGTAGACTGTGGCACTACGGCGTGTGATCCGAACGTGGATGGGTCCCGC 2374  
QY 2270 TGCAGGATGGCTGGATGGGGCAGCCTGCGACCGGGCTGCCACCCCGCTGTGCC 2329  
Db 2375 TGTGAAGAGGGCTGGACAGGCGCAGCTTGTGACCAAGCGCGTGTGCCACCCCGCTGCATT 2434

QY 2330 GAGCATGGGACCTGCCCGACGGCAAGTGGAGTGCAGCCCTGGCTGGAATGGCGAACAC 2389  
Db 2435 GAGCAGGGACCTGTAAAGATGGCAATGTGAATCCGAGAGGGGTGGAATGGTGAACAC 2494  
QY 2390 TGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAAT 2449  
Db 2495 TGCACCATTTGAT-----GGCTGCCCTGATTGTGCAAC 2527  
QY 2450 GGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCCGTGCCAGCTGGGCTGG 2509  
Db 2528 GGTAAACGGGAGATGCACACTGGGTCAAGAACAGTGGCAGTGTCTGCCAGACCGGCTGG 2587  
QY 2510 AGAGGAGCTGCTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAGACAATGAT 2569  
Db 2588 AGAGGGCTGGATGCAACGTTGCCATGGAAACCTCCTGCGCTGATAACAAGGATAATGAG 2647  
QY 2570 GGAGATGGCCTGGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATC 2629  
Db 2648 GGAGATGGCCTGGTGGACTGCCCTGGACCCCTGACTGCTGCCCTACAGTCAGCCTGTCAAG 2707  
QY 2630 AACCCGCTGTGCTGGTCCCTAACCCCTCTGACATCATCCAGGAGACACAGGTCCT 2689  
Db 2708 AGCCTGCTCTGCCGGGGTCTCGGAGCCCTTGGACATCATTCAGCAAGGTTCAGACAGAC 2767  
QY 2690 GTGTACAGCAGAAACCTACACTCCTTCTATGACCGCATCAAGTTCTCTGTGGCAGGGAC 2749  
Db 2768 TGGCCTGCAGTGA-----AGTCTTCTATGACCGCATCAAGCTCTTGGCAGGCAAGGAC 2821  
QY 2750 AGCAGGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGCACTGTGTGTTATTCGT 2809  
Db 2822 AGCAGCCACATCATCTCTGGAGACACCCCTTCAATAGCAGCCTGGTGTCTCTGATCCGA 2881  
QY 2810 GGCCAAGTGTGACATCAGATGGAACCCCTGTTGGTGTGAACATCAGTTTGTCAAT 2869  
Db 2882 GGCCAAGTAGTAACCATGGATGGGACTCCCTTGGTGGTGTGAATGTCTTTTGTCAAG 2941  
QY 2870 AACCTCTCTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTTGGTGACAAAT 2929  
Db 2942 TACCCAAAATATGGCTACACCATCACTGCCAGGATGGCAGCTTGACCTGATTGCCAAT 3001  
QY 2930 GGCGGCATCTCCATCATCTGCGGTTGAGCGGGCACCTTTTCATCACACAGGAGCACACC 2989  
Db 3002 GGGGGTCTGCCCTTGACTCTTCACTTTGAGCGAGCCCTTTTCATGAGCCAGGAGCGCAC 3061  
QY 2990 CTGTGGCTGCCATGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGGAG 3049  
Db 3062 GTGTGGCTGCCATGGAACAGCTTCTATGCCATGGACACCCCTGGTAATGAAGACCGAGGAA 3121  
QY 3050 AATGAGATTCCAGCTGTGACCTGAGCAATTTTGGCCGCCCCAATCCAGTCGTCTCTCA 3109  
Db 3122 AACTCCATCCCAGCTGTGACCTCAGTGGCTTTGTCCGGCCAGATCCAATCATCATCTCC 3181  
QY 3110 TCCCCACTGACGTCCTTCCCGAGCTCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGC 3169  
Db 3182 TCTCCTCTGTCCACCTTCTTCAGCGGCTTCCCTGCCCTCGAAACCCCATTTGTGCCTGAGACC 3241  
QY 3170 CAGGCTTTGAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGC 3229  
Db 3242 CAGGTTCTTCATGAAGAAATGAGCTCCCTGGTACCAATGTGAAGCTCCGTTATCTCAGC 3301  
QY 3230 AGCCGACCCCTGGCTACAAATCTGTCTGAGGATCAGCCTCACCCACCCGACCAATCCCC 3289  
Db 3302 TCTAGAACTGCAGGGTATAAGTCGCTGTGAAGATCACCATGACGAGTCCACAGTGGCC 3361  
QY 3290 TTCAACCTCATGAAGGTGCACCTCATGGTAGCGGTGGAGGCGCCCTCTTCAGGAAAGTGG 3349  
Db 3362 TTGAACCTCATCAGGGTTCACTTGATGTTGCTGTAGAGGGGCATCTCTTCCAGAGTCA 3421  
QY 3350 TTGCTGCAAGCCCGCAGACCTGTCTCTATTTATTTTCAATTTGGACAAGACAGACGCTTACAAC 3409  
Db 3422 TTCCAGGCTTCTCCCAACCTAGCTACACATTCATCTGGACAAGACAGATGCTTATGGC 3481  
QY 3410 CAGAAAGGTGTTGGGCTTTCAGAAAGCCTTTGTTTCCGTGGGTATGAATATGAATCCTGC 3469

Db 3482 CAAAGGGTTATGGCCTATCGGATGCTGTTGTGTTGTTGGTTTGAATATGAGACCTGC 3541  
QY 3470 CCAGATCTAATCCTGTGGGAAAAAAGAACAAACAGTGTGTCAGGGCTATGAATGACGCG 3529  
Db 3542 CCAGTCTCATCCTGTGGGAGAAAAGGACAGCCCTGCTTCAGGGATTGAGCTGGACCCCT 3601  
QY 3530 TCCAAGCTTGGAGGATGGAGCCTTAGACAAAACATCATGCCCTCAACATTCAAAAGTGGTATC 3589  
Db 3602 TCCAACCTTGGAGGCTGGTCCCTTGGACAAAACACACACCCCTCAATGTGAAAAGCGGAATA 3661  
QY 3590 CTGCACAAAGGGAATGGGAGAAACCCAGTTTGTGTCTCAGCAGCCCTCCTGTTCATTTGGGAGC 3649  
Db 3662 CTACACAAAGGACAGGGAGAACCCAGTTTCTGACCCAGCAGCCTGCCATCATCAGCAGC 3721  
QY 3650 ATCATGGGCAATGGGCGCGGAGAAAGCATCTCCTGCCCCAGCTGCAACGSCCTTGCTGAC 3709  
Db 3722 ATCATGGGCAACGGTCGCGCAGAAAGCATCTCCTGTCCAGCTGCAATGGCCTTGCTGAA 3781  
QY 3710 GGCAACAAGCTCCTGGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTG 3769  
Db 3782 GGCAACAACCTGTTAGCCCTGTGGCCCTGGCTGTGGGATCGATGGGAGCCTCTTTGTT 3841  
QY 3770 GGTGATTTCAACTACATTAGAAGGATCTTCCCCTCTGGAAATGTCAACCACATCCTAGAG 3829  
Db 3842 GGTGACTTCAACTATATCCGGCGCATCTTCCCCTCTCGAAATGTGACCACTATCTTGGAG 3901  
QY 3830 CTGAGGAATAAAGATTTACAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACA 3889  
Db 3902 TTACGAAATAAAGATTTAAACATAGCAACAGCCAGGACACAAAGTACTACTTGGCTGTG 3961  
QY 3890 GACCCCATGAGTGGGGCGCTCTTCCCTTCTGACAGCAACAGCCGGCGGTCTTTAAATC 3949  
Db 3962 GACCCCGTGACTGGCTCACTCTACGTCTCTGACACCAACAGTCGCCGAATCTACCGAGTC 4021  
QY 3950 AAGTCCACTGTGGTGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGT 4009  
Db 4022 AAGTCTCTGAGCGGAGCCAAAGACCTGGCTGGAATTCGGAAGTTGTGGCAGGACTGGC 4081  
QY 4010 GACCAGTGCCTCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCC 4069  
Db 4082 GAACAAATGTCTACCCCTTGATGAAGCCCGCTGTGGGGATGGAGGAAGGCTGTGGACGCC 4141  
QY 4070 ACACTCACCAATCCCAGGGTATTACAGTGGACAAAGTTTGGCTGATCTACTTCGTGGAT 4129  
Db 4142 ACCCTGATGAGCCCCAGAGTATTGCAGTAGACAAAGATGGGCTTATGTACTTTGTTGAT 4201  
QY 4130 GGCACCATGATCAGACGCATCGATCAGAATGGGATCATCTCCACCCCTGCTCGGCTCTAAT 4189  
Db 4202 GCCACCATGATCCGGAAGTGGACCAAAACGGAATCATCTCCACCCCTGCTGGGCTCCAAT 4261  
QY 4190 GATCTCACATCAGCCCGGCACTCAGCTGTGATTTCTGTTCATGGATATTTCCAGGTAAGA 4249  
Db 4262 GACCTCACAGCTGTCCGACCACTGAGCTGTGACTCGAGCATGGACGTGGGCCAGGTCCGT 4321  
QY 4250 CTGGAGTGGCCACAGACTTAGCCATCAACCCCAATGGACAACTCACTTTATGTCTCTCGAC 4309  
Db 4322 CTAGAATGGCCGACAGACCTCGCCGTCAACCCCATGGACAACTCCCTGTACGTTCTGGAG 4381  
QY 4310 AACAAATGTGTCCTGCAATCTCTGAAAAACCAAGGTGCGCATTTGTGCGCGGAGGCC 4369  
Db 4382 AACAAAGTCATCCTGCGGATCAGGAGAACCCAGGTGAGCATCATCGCGGACGGCCT 4441  
QY 4370 ATGCACTGCCAGGTCCCTGGCATTGACCACTTCCCTGCTAAGCAAGGTGGCCATCCACGCA 4429  
Db 4442 ATGCACTGCCAGGTTCGCGCATCGACTACT---CGCTCAGCAAACTCGCCATCCACTCT 4498  
QY 4430 ACCCTGGAGTCAGCCACCGCTTTGGCTGTTTCAACAAATGGGGTCCCTGTATATTGCTGAG 4489  
Db 4499 GCCTGGAATCAGCCAGCGCATTTGCCATTTCTCACACTGGGGTGTCTACATCACTGAG 4558  
QY 4490 ACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTACCACTAGTGGAGATCTCACTC 4549

Db 4559 ACGACGAGAGAAGATCAACCGCCTACGCCAAGTCACCACCAATGGAGAGATCTGCCTC 4618  
QY 4550 GTTGTGGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTTCT 4609  
Db 4619 TTAGCCGGGGCGCCTCAGACTGTGACTGCAAAAACGATGTCAACTGCACTCTGCTACTCG 4678  
QY 4610 GGAGACGATGGTTATGCCRAAGGATGCAAAAGTTAAATACCCCATCTTCCCTTGGCTGTGT 4669  
Db 4679 GGAGATGACGCTTACGCCACGGACGCCATCCTGAACTCGCCGTCTCCTTAGCCGTGGCT 4738  
QY 4670 GCTGATGGGAGCTCTACGTGGCCGACCTTTGGGAACATCCGAATTCGGTTTATCCGGAAG 4729  
Db 4739 CCGATGGCACCATCTACATTGCAGACCTTTGGGAATATCCGGATCAGGGCGGTGAGCAAA 4798  
QY 4730 AACAAAGCCTTTCCCTCAACACCCAGAAACATGTATGAGTGTCTTTCACCAATTGACAGGAG 4789  
Db 4799 AATAAAACCCGTTCTTAACCGATTCAACCAAGTATGAGGTGCATCTCCGGGAGAACAGGAA 4858  
QY 4790 CTCATCTGTTTGATACCAACCGGAAGCACCTGTACACCCAAAGCCTGCCACAGGAGAC 4849  
Db 4859 TTGTACGTGTTCAAACGCTGATGGTATCCATCAGTACACTGTGAGTCTGGTGACTGGGAG 4918  
QY 4850 TACTGTACAACTTCACCTACACTGGGACGGGACGCGACATCACACTCATCACAGACAAAT 4909  
Db 4919 TACTGTACAAATTCACATACAGCGCTGACAAATGACGTCAACCGAGTTGATTGACAAAC 4978  
QY 4910 GGCAACATGGTAAATGTCGCGGAGACTCTACTGGATGCCCCCTCTGGCTGGTGTCCCA 4969  
Db 4979 GGGAAATCCCTAAAGATCCGCGGGACAGCAGTGGCATGCCCCGCCACCTGCTCATGCCG 5038  
QY 4970 GATGGCACAGGTGTAAGTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACACA 5029  
Db 5039 GATAATCAGATTATCACCCCTACTGTGGCACCAATGGAGGCCTCAAAGCCGTGTCCTACT 5098  
QY 5030 CAAGGACACAGGTTGGCCATGATGACATACCATGGCAATTCGGCCTTCTTGGCAACCAAA 5089  
Db 5099 CAGAACCTGGAGCTGGSCCTCATGACTTATGATGGGAACACTGGACTCCTAGCCACCAAG 5158  
QY 5090 AGCAATGAAAAACGGATGGACAAACATTTTATGAGTACGACAGCTTTGGCCGCTGACAAAT 5149  
Db 5159 AGTATGAAACCGGATGGACAACTTTTATGACTATGACCAAGGGCCGTCTGACCAAT 5218  
QY 5150 GTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCAT 5209  
Db 5219 GTGACCCGCCCCACGGGCGTGGTGACCACTGTGCACCCGGGAAATGGAGAAATCTATCACC 5278  
QY 5210 GTCACGGTAGAGACCTCCAGCAAGGATGAT---GTACCATAAACCAACCAACCTGTCTGCC 5266  
Db 5279 ATTGACATTGAGAACTCCAAACGGGATGATGACGTCTACTGTGATCACCAACCTCTCCTCC 5338  
QY 5267 TCAGCGCCTTCTACACACTGTGCAAGACCAAGTCCGGAAACAGCTACTACATCGGGGCC 5326  
Db 5339 GTGGAGGCCTCCTATACAGTGGTACAAGATCAAGTGGCAAAACAGCTACCAAGCTCTGCAAT 5398  
QY 5327 GATGGCTCCTTGGGCTGCTGTGGCCAAACGGCATGGAGTGGCGCTGCAGACTGAGGCC 5386  
Db 5399 AATGGAACCTCGGGGTGATGTACGCCAAACGGCATGGGTGTTCAGCTTCCACAGTGAGGCC 5458  
QY 5387 CACTTGTGGCTGGCACCGTCAACCCCAACCGCTGGCAAGAGGAATGTACCGTGCCTATC 5446  
Db 5459 CACGTCTCGCAGGCACCATCACCCCAACCATCGGGCGCTGCAACATCTCTCTGCCCATG 5518  
QY 5447 GACAAACGCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGAGAGGAAATGTACCGTGCCTATC 5506  
Db 5519 GAGATGGCCTGAACTCCATCGAGTGGCGCCTGAGGAAGGAACAGATCAAAGGCAAGTGC 5578  
QY 5507 ACTGTCTTTGGGCGCGGCTGGGGTGCACAAACCGAAATCTCCTATCTCTGGACTTTGAT 5566  
Db 5579 ACCATCTTTGGGAGGAAGCTTCGGGTCCACGGAAGGAATCTCCTGTCCATTTGATTATGAC 5638  
QY 5567 CGCGTAACACGCACAGAGAAAGATCTATGATGACCAACCGCAAGTTTCAACCTTCGGATTCTG 5626  
Db 5639 CGAAATATCCGTACGGAGAAAGATCTACGATGACCACCGGAAATTCACCTTGAGGATCATC 5698



Db 7838 TCTAGCATCGCCAGTGAGGACAGCCGCAAGTAGCATCCGTGTTGAACAATGCCTACTAC 7897  
QY 7844 CTAGAGAACTGCACTTACCATTGATGGGTGGATACCCATTACTTTGTGAACACGGA 7903  
Db 7898 TTAGACAAGATGCACTACAGCATCGAGGGCAAGGACACACACTACTTTGTGAAGATCGGC 7957  
QY 7904 CTTTCAGAAAGTGACCTTGGCCCATCCTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAA 7963  
Db 7958 GCCGGGATGGTACCTGGTACGCTAGGAACCAACCATTTGGGCGCAGGTGCTGGAGAGT 8017  
QY 7964 GGGGTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTAC 8023  
Db 8018 GGGGTGAACGTGACCGTGTACAGCCCAACGCTGCTGCTGTAATGGCAGGACTCGAAGTTT 8077  
QY 8024 ACAGACATCCAGCTCAGTACGGGGCACTGTGCTTGAACACACGCTACGG-----G 8074  
Db 8078 ACCAACATTGAGTTCAGTACTCCACGCTGCTGCTCAGTATCCGCTACGGCCTCACCCCC 8137  
QY 8075 ACAACGTTGGATGAGGAGAAGGCACGGGTCTCGAGCTGGCCCGGCAAGAGCCGTGCGC 8134  
Db 8138 GACACGCTGGACGAGAAAGGCCCGCTCCTGGACCAAGCGGGACAGAGCCCTGGGT 8197  
QY 8135 CAAGCGTGGCCCGGACAGCAGAGACTGCGGGAAGGGAGGAAGGCTGCGGGCCTGG 8194  
Db 8198 ACTGCCTGGCCCAAGGACAGCAGAAAGCCAGGACGGGAGAGAGGGCAGCCCTGTGG 8257  
QY 8195 ACAGAGGGGAGAAAGCAGCAGGTGCTGAGCACAGGGCGGGTGCAAGGTACGACGGCTTT 8254  
Db 8258 ACGAGGGCGAGAAAGCAAACTCCTGAGCACGGGACGGGTACAAGGTATGAGGGCTAT 8317  
QY 8255 TTCGTGATCTCTGTCGAGCAGTACCCAGAACTGTGACAGACAGCGCCCAACAACATCCACTTC 8314  
Db 8318 TACGTACTTCGGTGGAAACAGTACCCGGAGCTGGCAGACAGTAGCAGCAACATCCAGTTC 8377  
QY 8315 ATGAGACAGAGCGAGATGGCCGGAGGTGACAGA 8348  
Db 8378 TTAAGACAGAATGAGATGGAAAGAGGTAACAAA 8411

RESULT 14

US-09-808-602-77  
; Sequence 77, Application US/09808602  
; Patent No. US20020155115A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herriman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-697 CIP  
; CURRENT APPLICATION NUMBER: US/09/808,602  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/800,198  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 77  
; LENGTH: 8797  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-808-602-77

Query Match 40.5%; Score 3384.8; DB 9; Length 8797;  
Best Local Similarity 64.8%; Pred. No. 0;  
Matches 5467; Conservative 0; Mismatches 2712; Indels 255; Gaps 20;

QY 35 ATGGACGTGAAGGAGAGAAAGCCTTACCGCTCGTGTACCCCGCGCGCGAGCGC 94  
Db 113 ATGGATGTAAAGGACCGGACATCGCTCTTTGACAGGGGACGGTGTGGCAAGAGTGT 172  
QY 95 CGCTACACAGCTCGTCCGCGGACAGCGAGGAGGGGCAAG--CCCGCAGAAATCGTAC 151  
Db 173 CGCTACACAGCTCCTCTCTGGACAGTGAGGACTGCGGTGTGCCCACTCAGAGTCTCTAC 232  
QY 152 AGCTCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGC 211  
Db 233 AGTTCAGTGAGACCTTGAAGGCTTATGACCATGACAGCAGAATGCATATGGAACCGA 292  
QY 212 GTCAAGGACATTTGTGCCCGCAGGAGGCCGAGAAATTTCTGCCGCACAGGTGCCAATTTCACC 271  
Db 293 GTACACAGACCTGGTGACCGGGAGTCCGATGAGTTTCTAGACAAGGGACAAATTCACC 352  
QY 272 CTGCGGAGCTGGGGCTGGAAGAAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGAC 331  
Db 353 CTGGCAGAAATTGGGAATCTGCGAG--CCCTCCCCACACCCGAAGTGGTTACTGTTCGAC 409  
QY 332 ATTGGCCTGCCCCAATGCGSCTACTCCATGGGGGTGGCTCTGTATCCGACATGGAGGCT 391  
Db 410 ATGGGTATCCTCCACACGAGGCTACTCCCTGAGCACTGGGTCTGTATGCAGACTCGGACAC 469  
QY 392 GACACGGTGTCTCCCTGAGCACCCCGTGCCTGTGTGGGCGGAGCACACGGTTCAGGG 451  
Db 470 GAGGAGGGATGTCTCCAGAACATGCCATCAGACTGTGGGACGAGGGATAAAATCCAGG 529  
QY 452 CGCAGCTCCTGCTGTCCAGCGGGGCCAATTCCAAATCTCACACTCACCGACACCGAGCAT 511  
Db 530 CGCAGCTCTGGCTGTCCAGCGCGAGAACTCGGCCCTTACTCTGACTGACTCTGACAAT 589  
QY 512 GAAACACTGAGACTGATCATCTCGGGCGGCC----TGAGAAACCAACGCGCGGTCCCGAC 567  
Db 590 GAAATAAATCGGATGACGACAATGGTCTGCCCATTTCCACCTACATCCTCGTCTAGCCTC 649  
QY 568 GCCGCGCGCGCGCTCTCGACGCGCCACACCCCAACAGCAGCACCAACGCGGCT----- 620  
Db 650 CTCCCATCTGCTCAGTGCCTAGCTCCCATAAATCCTCCACCAAGTTAGTCCAGATGCCA 709  
QY 621 ----- 620  
Db 710 TTGCTAGACAGCAACCTCTCCATCAGATCATGGACACCAACCTGTATGAGGAATTTCTCC 769  
QY 621 -CCATTAACTCCTTGAAACCGGGGCAACTTTCAGCGCGAGGAGCAACCCAGCCCGGCCCC 679  
Db 770 CCCAATTCATACCTGCTCAGAGCATGCTCAGGGCCCCAGCAAGCCTCCAGCAGTGGCCT 829  
QY 680 ACGGACCAAC-----TCGCTCTCCGAGAGCCCCCTGCCCCGGCGGCC 721  
Db 830 CCAACCAACACAGCCAGTCAACACTGAGGCCCTCTGTGCCACCCCTCATAAACCAACC 889  
QY 722 CAGGAGCCTGCCACGCCCGCAGAGAACTGGCTGTCTCAACAGCAACATCCCCCTGGAGACC 781  
Db 890 CTGTCCCAACCACTCCTCGGCCAACTCCCTCAACAGGAACTCACTGACCAATCGGCGG 949  
QY 782 AGGAACCTAGGCAAGCAGCCATTCTTAGGGACATTGACAGACAACCTC-----ATT 832  
Db 950 AGTCAAAATCCACGCCCGCAGCTCCTGGCCCCAACGACCTGGCCACCAACCCAGAGTCTGT 1009  
QY 833 GAGATGGACATTCTCGGGCGCTCCCGCCATGATGGGCTTACAGTGACGGGCACCTTCCTC 892  
Db 1010 CAGCTCCAGGATAGTGGGTGCTGAACAGTAACGTCCTCCACTGGAGACTCGGCACCTTCCT 1069  
QY 893 TTCAA---GCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCAGAGGTACCCACTG 949  
Db 1070 TTCAAAACGTCGTCGTGAAGCACACCCCTGTTTCAGCAGCTCTTCTCCGGGATACCCCTTG 1129  
QY 950 ACGTCCAGCACAGTGTACTCTCTCCGCCCGCCGACCCCTGCCCGCAGCACCTTCGCCCGG 1009  
Db 1130 ACCTCAGGACCGGTTTATACACCAACCAACCCCGCTGTCTGCCAGGAATACATTCTCCAGG 1189  
QY 1010 CCGGCTTTAACTCAAGAAAGCCCTCCAAGTACTGTAACTGGAAGTGCGCAGCCCTGAGC 1069

Db 1190 AAGGCCTTCAAGCTGAAGAAACCCCTCCAATACTGCAGTTGGAATGTGCTGCCCTGTCT 1249  
QY 1070 GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGCTGGCATACTTTGTGGCCATGCACCTG 1129  
Db 1250 GCCATCGCCCGCCCTCTCTTGGCCATTGTGTCGATATTTTCATAGCAATGCATCTG 1309  
QY 1130 TTTGGCCTAAACTGCGACCTGCGCCGATGGAGGGGCGAGATGTATGAGATCAGGGAGAC 1189  
Db 1310 CTCGGACTCAATTGGCAACTCCAGCCGGCAGATGGACACACCTTTAACA-----ATGGC 1363  
QY 1190 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCCTCCCTATACCCCTCAGGGGCGACTGGC 1249  
Db 1364 GTAAGGACCGGCTTACCAGGAAACGATGATGTGGCAACAGTGCCATCTGGAGGCAAAAGTG 1423  
QY 1250 TTAGAGACCCCTGACAGGAAAGGCAAGGAACCAAGAAAGCAAGAACCCAGTAGTTTCTTT 1309  
Db 1424 CCCTGGTCAATTGAAA----- 1439  
QY 1310 CCAGAGGACAGTTTCATAGATTCTGGAGAAATTGATGTGGAAAGCGGAGCCTCCAGAAAG 1369  
Db 1440 -----ACAGCAGCATAGACAGTGGCGAAGCAGAAGTTGGTCGGCGGTGACACAGGAA 1492  
QY 1370 ATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGTGAAA 1429  
Db 1493 GTCCACACAGGGGTGTTTGGAGGTCCAGATTTCACATCAGTCAGCCTCAATTCTTAAAG 1552  
QY 1430 TTCAATGTGTCTGGGAAAGGCGAGCCCTGGTTGGCATTTATGGCAGAAAAGCCTCCCT 1489  
Db 1553 TTCAACATCTCCCTGGCAAGGATGCCCTCTTCCGTTGTCTATATAAGGAGAGACTACCA 1612  
QY 1490 CTTTACATACACAGTTTGACTTTGTGAGCTGTCTGGATGCGAGGAGGCTCCTAACCCAG 1549  
Db 1613 CCGTCTCATGCCAGTATGACTTTCATGGAACGCTGGATGGA-----AG 1657  
QY 1550 GAGGCGGAGCTAGAGGGGACCCCGCCAGTCTCGGGGAACTGTGCCCCCTCCAGC 1609  
Db 1658 GAGAAATGGAGCGTGTGAGTCGCCCGAGGAAACGCCGAGGATCCAGACTCTGGTGCAG 1717  
QY 1610 CATGAGACAGGTTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTACAAAT 1669  
Db 1718 AACGAGGCTGTGTTGTGCACTACITGGATGTGGGCTGTGGCACTGGCCCTTCTACAAAT 1777  
QY 1670 GACGGAAAGGAGTCAGAAAGTGGTTTCTTCTCACCACCTGCCATTGAGTCGGTGGATAAC 1729  
Db 1778 GACGGCAAGGACAAGGAGATGGTCTCCTTCAACACTGTTGTCTTAGATTTCAGTGCAGGAC 1837  
QY 1730 TGCCCCAGCAACTGCTATGGCAATGGTGACTGCATCTCTGGACCTGCCACTGCTTCTCTG 1789  
Db 1838 TGTCCACGGAACACTGTACCGGAAACGGTGAATGCGTGTCTGCACTGTGTACTGTTCCCA 1897  
QY 1790 GGTTCCTGGCCCCGACTGTGGCAGAGCCTCTGCCCCGCTCTGTAGCGGAAATGGC 1849  
Db 1898 GGATTCCTAGTGCAGACTGTGCTTAAAGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1957  
QY 1850 CAATACATGAAGGCAGATGCTTGTGCCACAGTGGCTGGAAAGCGCTGAGTGGCATGTG 1909  
Db 1958 CAGTATTCTAAAGGAACGTGCCAGTGTACAGCGGCTGGAAAGGTGCAGAGTGTGATGTG 2017  
QY 1910 CCCACCAACAGTGTATCGATGTGGCTGCAGCAACCAAGGACCTGCATCAGCGGCACC 1969  
Db 2018 CCTATGAACCAATGTATCGATCCTTCTGTGGGGCCATGGCTCCTGCTGCTGCTGCTGCTG 2077  
QY 1970 TGCACTGTCAACCTGGCTACAAGGCGGAGAGCTGTGAGGAAGTGGACTGCATGGAACCCC 2029  
Db 2078 TGCCTGTGTGCTGTGGCTACAAGGCGGAGCACTGTGAGGAAGTTGATTGCTTGGATCCT 2137  
QY 2030 ACATGTTGAGGCGGGGTGTCTGCGTGAAGGCGGAATGCCATTGCTTTGTGGGATGGGA 2089  
Db 2138 ACCTGCTCAGCCATGGTGTCTGTGTGAATGGAGAGTGTCTATGCAGCCCCGCTGGGT 2197  
QY 2090 GGCACCAACTGCGAGACCCCGGCGCCACATGCTTAGACAGTGTTCAGGCCACGGAACC 2149

Db 2198 GGTCTCAACTGTGAGCTGGCGAGGGTCCAGTGCCCGAGACCAGTGTAGTGGCATGGCACT 2257  
QY 2150 TTCTCCCGGACACCGGGCTTTTGAGCTGTGACCCCAAGCTGGACTGGACACGACTGTTCT 2209  
Db 2258 TACCTCCCTGACTCCGGCTCTGCGCTGTGATCCGAACCTGATGGGTCCCGACTGCTCT 2317  
QY 2210 ATCGAGATCTGTGCTCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGCGACCTGCCGC 2269  
Db 2318 GTTG--TGTGCTCAGTAGACTGTGGCACTACGGCGTCTGCACTCGGGGAGCCTGCCGC 2374  
QY 2270 TCGAGGATGCTGTGGATGGGGGCGACCTTCGACACGCGGGCCTGCCACCCGCTGTGCC 2329  
Db 2375 TGTGAAGAGGCTGGACAGGCGCAGCTTGTGACCAAGCGCTGTGCCACCCCGCTGCATT 2434  
QY 2330 GAGCATGGGACCTCCCGCAGCGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACAC 2389  
Db 2435 GAGCACGGGACCTGTAAAGATGGCAAAATGTGAATCCGAGAGGGCTGGAATGGTGAACAC 2494  
QY 2390 TGCACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTCGCCCTGGTTGTGCAAT 2449  
Db 2495 TGCACCATTTGAT-----GGCTGCCCTGATTTGTGCAAC 2527  
QY 2450 GGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTG 2509  
Db 2528 GGTAAACGGGAGATGCACACTGGGTGAGAAACAGCTGGCAGTGTGTCTGCCAGACCGGCTG 2587  
QY 2510 AGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCTTCCGCTGCAAGCAAGACAAATGAT 2569  
Db 2588 AGAGGGCTGGATGCAACGTTGCCATGGAACCTCCTGCGCTGATAACAAGGATAATGAG 2647  
QY 2570 GGAGATGGCTGGTGGACTGCATGGACCTGACTGCTGCTCCAGCCCTGTGCCATATC 2629  
Db 2648 GGAGATGGCTGGTGGACTGCCCTGGACCTGACTGCTGCTTACAGTCAGCCTGTGAGAAC 2707  
QY 2630 AACCCGCTGTGCTTGGCTCCCTAAACCTCTTGACATCATCCAGGAGACACAGGTCCCT 2689  
Db 2708 AGCCTGCTGCTCGGGGTCTCGGACCCCTTGACATCATTCAGCAAGGTGAGACAGAC 2767  
QY 2690 GTGTACAGCAGAAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGCAGGAC 2749  
Db 2768 TGGCCTGCAGTGA-----AGTCTTCTATGACCGCATCAAGCTCTTGGCAGGCAAGGAC 2821  
QY 2750 AGCAGCGCAGATAATCCCGGGGAGAACCCCTTTGATGGAGGCGCATGCTTGTGTTATCGT 2809  
Db 2822 AGCAGCCACATCATCTCTGGAGACAACCCCTTCAATAGCAGCTGGTGTCTCTGATCCGA 2881  
QY 2810 GGCCAAGTGTGACATCAGATGGAACCCCTTGGTTGGTGTGAACATCAGTTTGTGCAAT 2869  
Db 2882 GGCCAAGTAGTAACCATGGATGGGACTCCCTTGGTGGTGTGAATGTGTCTTTTGTCAAG 2941  
QY 2870 AACCTCTCTTTGGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAAT 2929  
Db 2942 TACCCAAAATATGGCTACACCATCACTCGCCAGGATGGCAGCTTTGACCTGATTGCCAAT 3001  
QY 2930 GCGGCGCATCTCCATCATCTGCGGTTGAGCGGGCACCTTTTCATCACACAGGAGCACACC 2989  
Db 3002 GGGGGTTCTGCCCTTGACTCTTCACTTTGAGCGAGCCCCCTTTCATGAGCCAGGAGCGACA 3061  
QY 2990 CTGTGCTGCCATGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGGAG 3049  
Db 3062 GTGTGGCTGCCATGGAAACAGCTTCTATGCCATGGACACCCCTGGTAATGAAGACCGAGGAA 3121  
QY 3050 AATGAGATCCCAGCTGTGACCTGAGCAATTTTTCGCCGCCCAACCCAGTCGCTCTCTCCA 3109  
Db 3122 AACTCCATCCCCAGCTGTGACCTGAGTGGCTTTGTCCGCCAGATCCAATCATCATCTCC 3181  
QY 3110 TCCCACTGACGCTCCTTCGCCAGCTCCTGTGAGAGAAAGGCCCCCATTTGTGCCGGAATTT 3169  
Db 3182 TCTCCTCTGTCCACCTTCTTCCAGCGCTTCCCTGCTCGAATCCCATTTGTGCTGAGACC 3241  
QY 3170 CAGGCTTTCAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTAGCTAGC 3229  
Db 3242 CAGGTTCTTCATGAAGAAATGAGCTCCCTGGTACCAATGTGAAGCTCCGTTATCTCAGC 3301

QY 3230 AGCGGACCCCTGGCTACAAATCTGTCTGAGGATCAGCCTCACCCACCGACCATCCCC 3289  
Db 3302 TCTAGAACTGCAGGGTATAAGTCGCTGCTGAAGATCACCATGACGCAGTCCACAGTGCCC 3361  
QY 3290 TTCAACCTCATGAAGTGACACCTCATGGTAGCGGTGGAGGCGCGCTCTTCAGGAAGTGG 3349  
Db 3362 TTGAACCTCATCAGGGTTCACCTTGATGGTTGCTGTAGAGGGGCATCTCTCCAGAAGTCA 3421  
QY 3350 TTCGCTGCAGCCCCAGACCTGTCTCTATTATTTTCCTGGGTATGAATATGAATCCTGC 3409  
Db 3422 TTCAGGCTTCTCCCAACCTAGCCTACACATTTCATCTGGACAAGACAGATGCTTATGGC 3481  
QY 3410 CAGAAGGTGTTTGGGCTTTCAGAAGCCTTTGTTTCCTGGGTATGAATATGAATCCTGC 3469  
Db 3482 CAAAGGTTTATGGCCTATCGGATGCTGTGTGTCGTGTTGGTTTGAATATGAGACCTGC 3541  
QY 3470 CCAGATCTAATCCTGTGGGAAAAAGAACACAGTGTCTGACGGGCTATGAATTTGACGG 3529  
Db 3542 CCCAGTCTCATCCTGTGGGAGAAAGGACAGCCCTGCTTCAGGGATTTCAGCTGGACCCCT 3601  
QY 3530 TCCAAGCTTGGAGGATGGAGCCTTAGACAAACATCATGCCCCCTCAACATTCAAAGTGGTATC 3589  
Db 3602 TCCAACCTTGGAGGCTGGTCCCTGGACAAACACACACCCCTCAATGTGAAAAGCGGAATA 3661  
QY 3590 CTGCACAAAGGGAATGGGCGCGGAGAACAGATTGTTGTCTCAGCAGCCTCCTGTCTCATTTGGAGC 3649  
Db 3662 CTACACAAAGGACAGGGAGAACCAAGTTCCCTGACCCAGCAGCCTGCCATCATCACGAGC 3721  
QY 3650 ATCATGGGCAATGGGCGCGGAGAACAGCATCTCCTGCCCCAGCTGCAACGGGCTTGTCTGAC 3709  
Db 3722 ATCATGGGCAACGGTCGCGCAGAGAACAGCATCTCCTGTGCCAGTGCATGGGCTTGTCTGAA 3781  
QY 3710 GGGCAACAAGCTCCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTG 3769  
Db 3782 GGGCAACAACCTGTTAGCCCTGTGGCCCTGGCTGTGGGATCGATGGGAGCCTCTTTGTT 3841  
QY 3770 GGTGATTTCAACTACATTAGAAGGATCTTCCCTCTGGAAATGTCAACCAATCCTAGAG 3829  
Db 3842 GGTGACTTCAACTATATCCGGCGCATCTTCCCTCTCGAAATGTGACCAATCTTTGGAG 3901  
QY 3830 CTGAGGAATAAAGATTTTCAGACATAGTCACAGTCCAGCACACAAATACTACTCTGGCCACA 3889  
Db 3902 TTACGAATAAAGAGTTTAAACATAGCAACAGCCAGGACACAAAGTACTACTTGGCTGTG 3961  
QY 3890 GACCCCATGAGTGGGCGGCTCTCTCTTCTGACAGCAACAGCCGGCGGTCTTTAAATC 3949  
Db 3962 GACCCCGTGACTGGCTCACTCTACGTCTCTGACACCAACAGTCGCCGAATCTACCGAGTC 4021  
QY 3950 AAGTCCACTGTGGTGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTGGGGGACAGGT 4009  
Db 4022 AAGTCTCTGAGCGGAGCCAAAGACCTGGCTGGAATTCGGAAGTTGTGGCAGGACTGGC 4081  
QY 4010 GACCAAGTGCCTCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCC 4069  
Db 4082 GAACAAATGTCTACCCCTTTGATGAAGCCCGCTGTGGGGATGGAGGAAGGCTCTGTGGACGCC 4141  
QY 4070 ACACTCAACCAATCCCAGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGAT 4129  
Db 4142 ACCCTGATGAGCCCCAGAGGTATTGCAGTAGACAAAGATGGGCTTATGTACTTTGTTGAT 4201  
QY 4130 GGCACCATGATCAGACGCAATCGATCAGAATGGGATCATCTCCACCCTGCTCGGCTCTAAT 4189  
Db 4202 GCCACCATGATCCGGAAGGTGGACCAAAACGGAATCATCTCCACCCTGCTGGGCTCCAAT 4261  
QY 4190 GATCTCACATCAGCCCCGCACTCAGCTGTGATTTCTCATGGATATTTCCAGGTAAGA 4249  
Db 4262 GACCTCACAGCTGTCCGACCACTGAGCTGTGACTCGAGCATGGACGTGGCCAGGTCGGT 4321  
QY 4250 CTGAGTGGCCACAGACTTAGCCATCAACCCCAATGGACAACCTCACTTTATGTCTCTCGAC 4309  
Db 4322 CTAGAATGGCCGACAGACCTCGCCGTCAACCCCAATGGACAACCTCCCTGTACGTCTCTGGAG 4381

QY 4310 AACAAATGTGGTCTCTGCAAAATCTCTGAAAAACCACCAGGTGGGCATTTGTGCCCCGGAGGCC 4369  
Db 4382 AACAAAGTTCATCTCTGGGATCACGGAGAACCCACCAGGTACGATCATCGGGGACGGCT 4441  
QY 4370 ATGCACTGCCAGGTCCCTGGCATTGACCACTTCTCTGCTAAAGCAAGGTGGCCATCCACGCA 4429  
Db 4442 ATGCACTGCCAGGTTCCTGGCATTCGACTACT---CGCTCAGCAAACTCGCCATCCACTCT 4498  
QY 4430 ACCCTGGAGTCAGCCACCGCTTTGGCTGTTTTCACAAATGGGGTCTCTGTATATTGCTGAG 4489  
Db 4499 GCGTGGAAATCAGCCAGCGCCATGCCATTTCTCACACTGGGTGCTCTACATCATCTGAG 4558  
QY 4490 ACTGATGAGAAAAAGATCAACCGCATCAGGCAGGTACCACTAGTGGAGAGATCTCACTC 4549  
Db 4559 ACGGACGAGAAAGATCAACCGCCTACGCCAAGTCAACCAATGGAGAGATCTGCTC 4618  
QY 4550 GTTGTGGGGCCCCCAGTGGGTGTGACTGTAAAAATGATGCCAACTGTGATTTGTTTCT 4609  
Db 4619 TTAGCCGGGGCGGCTCAGACTGTGACTGCAAAAACGATGTCAACTGCACTGCTACTCG 4678  
QY 4610 GGAGACGATGTTATGCCAAGSATGCAAAAGTTAAATACCCCATCTTCCCTTGGCTGTGT 4669  
Db 4679 GGAGATGACGCTTACGCCACGACGCGCATCCTGAACTCGCGCTCTCCTTAGCCGTGGCT 4738  
QY 4670 GCTGATGGGAGCTCTACGTGGCGGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAG 4729  
Db 4739 CCGGATGGCACCATCTACATTCAGACCTTGGGAATATCCGGATCAGGGCGGTGAGCAAA 4798  
QY 4730 AACAGCCTTTCTCAACACCCAGAAACATGTATGAGCTGTCTTCAACCAATTGACCAAGAG 4789  
Db 4799 AATAAACCCGTTCTTAACGCATTCAACCAAGTATGAGGTGCACTCCGGGAGAACAGGAA 4858  
QY 4790 CTCATCTGTTTGATACCCCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGAC 4849  
Db 4859 TTGTACGTGTTCAACGCTGTGTTATCCATCAGTACACTGTGAGTCTGGTACTGGGGAG 4918  
QY 4850 TACCTGTACAACTTCACTACTGGGACGGGACATCACACTCATCACAGACAACAAT 4909  
Db 4919 TACTGTACAAATTTACATACAGCGCTGACAATGACGTACCGAGTTGATTGACAACAAC 4978  
QY 4910 GGCAACATGGTAAATGTCCCGGAGACTCTACTGGGATGCCCTCTGGCTGTGGTCTCCA 4969  
Db 4979 GGGAAATTCCTAAAGATCCGCGGGACAGCAGTGGCATGCCCCGCCACCTGCTCATGCCG 5038  
QY 4970 GATGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACACACA 5029  
Db 5039 GATAATCAGATTATCACCCCTTACTGTGGGCACCAATGGAGGCTTCAAAGCCGTGTCCACT 5098  
QY 5030 CAAGGACACAGATTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAA 5089  
Db 5099 CAGAACTGGAGCTGGGCTCATGACTTATGATGGGAACACTGGACTCCTAGCCACCAAG 5158  
QY 5090 AGCAATGAAAAACGGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTGACAAAT 5149  
Db 5159 AGTGATGAAACCGGATGGACAACCTTTTATGACTATGACACAGAGGCGCTGTGACCAAT 5218  
QY 5150 GTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCAT 5209  
Db 5219 GTGACCCGCCCCACGGCGCTGTGACCAAGTCTGCACCGGAAATGGAGAAATCTATCACC 5278  
QY 5210 GTCCAGGTAGAGACCTCCAGCAAGGATGAT--GTCACCATAACCAACCTGTCTGCC 5266  
Db 5279 ATTGACATTGAGAACTCCAACCGGGATGATGACGTCACTGTGATCACCAACCTCTCCTCC 5338  
QY 5267 TCAGGCGCTTCTACACACTGCTGCAAGACCAAGTCCCGGAACAGCTACTACATCGGGGCC 5326  
Db 5339 GTGGAGGCTCCTATACAGTGTGTAAGATCAAGTGCGAACAGCTACCAGCTCTGCAAT 5398  
QY 5327 GATGGCTCCTTGGCGCTGCTGGGCCAACGGCATGGAGGTGGCGCTGCAGACTGAGCCC 5386  
Db 5399 AATGGAAACCTGCGGGTGTGTACGCCAACGGCATGGCTGTGAGCTTCCACAGTGGGCC 5458  
QY 5387 CACTTGTGGCTGGCACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACGCTGCCCATC 5446

Db 5459 CACGTCTCGCAGGACCATACACCCACCACCATCGGGCGCTGCAACATCTCTCTGCCCATG 5518  
 QY 5447 GACAAAGCCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGTC 5506  
 Db 5519 GAGAAATGGCCTGAACCTCCATCGAGTGGCGCTGAGGAAGAAACAGATCAAAGGCAAGTC 5578  
 QY 5507 ACTGTCTTTGGCGCCGGCTGGCGGTGCAACCGAAATCTCTATCTCTTGACTTTGAT 5566  
 Db 5579 ACCATCTTTGGGAAGAACTTCGGGTCCACGGAAGGAATCTCTCTCCATTTGATTAGAC 5638  
 QY 5567 CGCGTAACACGACAGAGAAGATCTATGATGACCAACCGAAAGTTCAACCTTCGGATTCTG 5626  
 Db 5639 CGAAATATCCGTACGGAGAAGATCTACGATGACCAACCGAAATTCACCTTGAGGATCATC 5698  
 QY 5627 TACGACAGGCGGGCGCCAGCCTCTGTGTCAACCGAGCAGAGGCTGAATGGTGTCAAC 5686  
 Db 5699 TATGACCAGGTGGCGCCGCCCTTCTCTGTGGCTCCGAGCAGTGGGCTGGCAGCCGTCAAT 5758  
 QY 5687 GTGACATACTCCCTGGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGA 5746  
 Db 5759 GTCTCTACTTCTCAATGGGCGCTTGGCGCGCTCCGAGCAGTGGGCTGGCAGCGAGAGG 5818  
 QY 5747 ATGGAATACGACAGGCGGGCGCCGATCATCATCCAGGATCTTCGCTGATGGGAAGACATGG 5806  
 Db 5819 ACAGACATTTGACAAGCAAGCGCGGATCGTGTCCCGCATGTTCCGCGACGGGAAGTCTGG 5878  
 QY 5807 AGCTACACATACTTAGAGAAGTCCATGGTGTCTACTACACAGCAGAGGCGAGTATATC 5866  
 Db 5879 AGTTATTCCTATCTTGACAAGTCCATGGTCTTCTGTCTACAGAGCCAAACGTGAGTACATA 5938  
 QY 5867 TTTGAGTTCGACAAGAATGACCGCCCTCTCTGTGACGATGCCCAACCGTGGCGCGGCGAG 5926  
 Db 5939 TTTGAATATGACTCTCCGATCGCCTCCAGCGAGTCACTATGCCAGTGTGCGCCCGGCAC 5998  
 QY 5927 AACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGCG 5986  
 Db 5999 AGCATGTCCACGACACCTCCATTTGGTTACATCCGAAACATTTACAACCCACCCGAAGC 6058  
 QY 5987 AATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCACTCTCTTCAACCTTCTACTCTG 6046  
 Db 6059 AATGCATCGGTCTTTTGACTACAGTATGACGGCCGCTCTCTAAAGACATCTTTCTTG 6118  
 QY 6047 GGCACCTGGCGAGGGTGATATACAAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTC 6106  
 Db 6119 GGCACCTGGCGCAGGTGTTCTACAAGTATGGAAACTCTCCAAGTTATCAGAGATAGTC 6178  
 QY 6107 TATGACACCAAGGTGATTTCACTATGAGGATGGGCACTCTCTTCAACCTTCTACTCTG 6166  
 Db 6179 TACGACAGCAGCCGTCACCTTTGGGTATGACGAGACCAACCGGTGTCCTGAAGATGGTC 6238  
 QY 6167 AACCTACAGAAATGAGGGCTTCACTGAGGAAGGCATGGTCAACGCGCCGTTTGTGACTACAAC 6226  
 Db 6239 AATCTCCAAAGTGGGGCTTCTCTGTACCATCAGGTACCGAAAGGTGGGCGCCCTGTG 6298  
 QY 6227 GACCGACAGATCTTCGCTTCACTGAGGAAGGCATGGTCAACGCGCCGTTTGTGACTACAAC 6286  
 Db 6299 GACAAGCAGATTTACAGGTTCTCTGAGGAAGGAATGATCAACGCGCAGGTTTGAATTATACC 6358  
 QY 6287 TAT--GACACAGCTTCCGGGTGACCGAGCATGAGGCTGTGATCAACGAGACCCCACTG 6343  
 Db 6359 TATCAGCAATAGCTTCCGATTTCCGATTCAGCATCAACCCGTCATTAGCGAGACTCCCTT 6418  
 QY 6344 CCCATTGATCTCTATCGCTATGATGATGTGTCAGGCAAGACAGACAGCTTTGGGAAGTTT 6403  
 Db 6419 CCTGTTGACCTTACCGCTATGACGAGATTTCCGGCAAGGTGGAACACTTCGGCAAGTTT 6478  
 QY 6404 GGTGTCTATTTACTATGACATTAACCAAGATCATCACCAAGCTGTCTATGACCCACCAAG 6463  
 Db 6479 GGGGTCTATCTACGACATCAACCAAGATCATCACCACTGCCCTCATGACGCTTAGCAAG 6538  
 QY 6464 CATTTTGTATGATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATG 6523

Db 6539 CACTTTGACACCCATGGGCGCATCAAGGAAGTGCAATATGAGATGTTCCGGTCCCTCATG 6598  
 QY 6524 TACTGGATGACCGTCCAGTATGATAACATGGGCGGAGTAGTGAAGAAGAGCTGAAGGTA 6583  
 Db 6599 TACTGGATGACTGTGCAATATGACAGTATGGGTAGGGTCAATCAAGAGGAACTGAAACTA 6658  
 QY 6584 GGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGATGCTGACGGCCAGCTGCAG 6643  
 Db 6659 GGGCCCTATGCCAAACACCAAAAGTACACCTATGACTATGACGGGACGGCCAGCTCCAG 6718  
 QY 6644 ACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTG 6703  
 Db 6719 AGTGTGGCGCTCAATGACCGGCTACCTGGCGCTATAGCTATGACCTCAATGGGAACCTG 6778  
 QY 6704 CACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGAC 6763  
 Db 6779 CACCTTCTAAACCCAGGAACAGTGTCTGCTCTATGCCCTTACGCTATGACCTCCGTGAC 6838  
 QY 6764 CGCATCACTCGGCTGGGTGACGTGCAATACAAAGTGGATGGATGGGCTTCTCTGAGGCG 6823  
 Db 6839 CGGATAACCAAGCTAGGGACGTGCAATACAAATCGATGACGATGGCTATTTGTGCCAG 6898  
 QY 6824 CGGGGGGTGATATCTTTGAGTACAACCTGAGCTGGCCTGCTCATCAAGGCCTACAACCGG 6883  
 Db 6899 AGAGGTGACACATCTTTGAATACAACTCAAGGGCTTCTGACGAGAGCATACAACAAG 6958  
 QY 6884 GCTGGCAGCTGAGTGTGAGTACCGGTACCGCTACGATGGCCTGGGCGCGCGTGTCCAGCAAG 6943  
 Db 6959 GCCAGCGGATGGAGCGTGCAGTACCGCTATGACGGAGTGGGCGCGGGCTTCTCTACAAG 7018  
 QY 6944 AGCAGCCACAGCCACCACTGAGTTCCTTCTATGCAGACCTGACCAACCCCAAGGTC 7003  
 Db 7019 ACCAACCTGGGCGCACCACTACAGTACTTCTACTCCGACCTCCACAACCCCAACAGTATC 7078  
 QY 7004 ACCACCTGTACAAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGA 7063  
 Db 7079 ACCCATGTTTACAACCACTCCAACCTGAGATCACCTCGCTCTACTATGACCTCCAGGCG 7138  
 QY 7064 CACCTCTTTGCCATGGAGCTGAGCAGTGGTGTGATGATTTTACATAGCTTTGTGACAAATC 7123  
 Db 7139 CACCTATTTGCCATGGAGAGCAGTGTGTTGAAGTATCTATGTCGCTCAGACAAACAG 7198  
 QY 7124 GGGACCCCTCTTGTCTCTTTAGTGAACAGGTTTGTGATGATCAAGCAAAATCCTGTACACA 7183  
 Db 7199 GGGACCCCTCTGGCTGTGACAGTATCAATGGCTCATGATCAAGCAAACTGACGTACACA 7258  
 QY 7184 GCCTATGGGAGATCTACATGGATACCAACCCCAACTTTTCAGATCATATAGGCTACCAT 7243  
 Db 7259 GCCTATGGGAGATCTACTATGACTCCAATCCAGACTTCCAGATGGTCTATGGCTTCCAC 7318  
 QY 7244 GGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGCGGAGATTATGATGTG 7303  
 Db 7319 GGAGGCCTCTATGACCCCTCACCAGCTCGTCCACTTTACTCAACGTGATTATGACGTG 7378  
 QY 7304 CTGGCCGACGCTGGACTAGCCCAAGACCAAGCTGTGGAAGCACTTTAGTAGCAGCAAC 7363  
 Db 7379 CTGGCAGGACGGTGGACGTCCCCCGACTACACCATGTGGAGGAACGT--GGGCAAGGAG 7435  
 QY 7364 GTCATGCCCTTTAATCTCTATATGTTCAAAAACAAACCCCATCAGCAACTCCAGGAC 7423  
 Db 7436 CCAGCCCTTCAACCTGTACATGTTCAAGAAACAACAATCTCTGAGCAATGAGCTGGAC 7495  
 QY 7424 ATCAAGTGTCTCATGACAGATGTTAAACAGCTGGCTGCTACCTTTGGATTCCAGCTACAC 7483  
 Db 7496 TTAAGAACTACGTGACAGACGTGAAGAGCTGGCTTGTGATGTTTGGATTTCAGCTCAGC 7555  
 QY 7484 AACGTGATCCCTGGTTATCCCAAACAGACATGGATGCCATGGAACCCCTCCTACGAGTC 7543  
 Db 7556 AACATCATCTCTGGATTCCCGAGAGCCAAAATGTATTTTGTGCTCCCCCTATGA--- 7611  
 QY 7544 ATCCACACACAGATGAAAAACGAGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAG 7603  
 Db 7612 -----ACTGTGAGAGTCAAGCAAGCGAGAACGAGAGCTCATTACAGGTGTCCAG 7663



Db 1010 CAGCTCCAGGATAGCTGGGTGCTGAACAGTAACAGTCCACCTGGAGACTCGGCACATTCCCTT 1069  
QY 893 TTCAA---GCCTGAGGCACCTCCCGCTCTCTCTGTCACACACATACACAGGGTACCCACTG 949  
Db 1070 TTCAAAACGCTGCTGGAAGCACACCCCTGTTTCAGCAGCTCTTCTCCGGGATACCCCTTG 1129  
QY 950 ACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCAGCACCTTCGCCCGG 1009  
Db 1130 ACCTCAGGACCGTTTATACACACACCCCGCTGCTGCCACCGAATACATTCTCCAGG 1189  
QY 1010 CCGGCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACCTGGAAGTGCAGCCCTGAGC 1069  
Db 1190 AAGGCCTTCAAGCTGAAGAAACCTCCAAATACTGCAGTTGGAATGTGCTGCCCTGTCT 1249  
QY 1070 GCCATCGTCATCTCAGCCACTCTGGTCACTCTGCTGGCATACTTTGTGGCCATGCACCTG 1129  
Db 1250 GCCATCGCCCGCCCTCTCTTGGCCATTGCTGGCATAATTTTCATAGCAATGCATCTG 1309  
QY 1130 TTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTATGAGATCACGGAGGAC 1189  
Db 1310 CTCGGACTCAATTGGCAACTCCAGCCGGCAGATGGACACACCTTTAACA-----ATGGC 1363  
QY 1190 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCCTCAGGGGGCACTGGC 1249  
Db 1364 GTAAGGACCGCTTACCAGGAAACGATGATGTGGCAACAGTGCCTATCTGGAGGCAAGTG 1423  
QY 1250 TTAGAGACCCCTGACAGGAAAGGCAAGGAAACCAAGAAAGGAAAGCCAGTAGTTTCTTT 1309  
Db 1424 CCCTGGTCAATTGAAA----- 1439  
QY 1310 CCAGAGGACAGTTTCATAGATTCTGGAGAAATTGATGTGGGAAGCGGAGCCCTCCAGAG 1369  
Db 1440 -----ACAGCAGCATAGACAGTGGCGAAGCAGAAAGTTGGTCCGCGGTGACACAGGAA 1492  
QY 1370 ATTCTCTCTGGCACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCACTGAAA 1429  
Db 1493 GTCCACACAGGGGTGTTTGGAGGTCCTCAGATTTCATCAGTCAGCCCTCAATTCTTAAAG 1552  
QY 1430 TTCAATGTGTCTTGGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAAGGCCTCCCT 1489  
Db 1553 TTCAACATCTCCCTGGGCAAGGATGCCCCTCTTCGGTGTCTATATAAGGAGAGGACTACCA 1612  
QY 1490 CTTTACATACAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCCTAACCCAG 1549  
Db 1613 CCGTCTCATGCCAGTATGACTTCATGGAACGCCTGGATGGAA-----AG 1657  
QY 1550 GAGGCGCGGAGCCTAGAGGGGACCCCGCGCAGTCTCGGGAACTGTGCCCCCTCCAGC 1609  
Db 1658 GAGAAAATGAGCGTGGTTCGAGTCGCCCGGAGGAAAGCCCGAGCATCCAGACTCTGGTGCA 1717  
QY 1610 CATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTGGCTTTTACAAAT 1669  
Db 1718 AACGAGGCTGTGTTGTGCAGTACTTGGATGTGGCCTGTGGCACTGGCCCTTCTACAAAT 1777  
QY 1670 GACGGAAAGGAGTCAGAAAGTGGTTTCTTCTTCCCACTGCCACTGCCATTGAGTCGGTGGATAAC 1729  
Db 1778 GACGGCAAGGACAGGAGATGGTCTCTCTCAACACTGTTGTCTTAGATTTCAGTGCAGGAC 1837  
QY 1730 TGCCCCAGCAACTGCTATAGGCAATGTTGACTGCATCTCTGGGACCTGCCACTGCTTCTCTG 1789  
Db 1838 TGTCCACGGAATGTCACGGGAACGGTGAATGCGTGTCTGGACTGTGTCACTGTTTCCCA 1897  
QY 1790 GGTTCCTGGGCCCCGACTGTGGCAGAGCCTCTGCCCCCGTCTCTGTAGCGGAAATGGC 1849  
Db 1898 GGATTCTAGGTGCAGACTGTGCTAAAGCTGCCCTGCCCTGTACTGTGCAGCGGAAATGGA 1957  
QY 1850 CAATACATGAAGGCAGATGCTTGTGCCACAGTGGCTGGAAAGCGCTGAGTGGCATGTG 1909  
Db 1958 CAGTATTCTAAAGGAAAGTGGCCAGTGTCTACAGCGGCTGGAAAGGTGCAGAGTGTGATGTG 2017  
QY 1910 CCCACCAACCACTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCAACGGGCACC 1969

Db 2018 CCTATGAACCAATGTATCGATCCTTCTCTGGGGCCATGGCTCCTGCATTGATGGGAAC 2077  
QY 1970 TGCATCTGCARACCTTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCC 2029  
Db 2078 TGCCTGTGTGCTGCTGCTACAAGGGCGAGCACTGTGAGGAAGTGTGCTTGGATCCT 2137  
QY 2030 ACATGTTCAGGCCGGGTGTCTGCTGAGAGGCGAATGCCATTGCTTTGTGGATGGGA 2089  
Db 2138 ACCTGCTCCAGCCATGTTGCTGTGTAATGGAGAGTGTCTATGAGCCCCCGCTGGGT 2197  
QY 2090 GGCACCAACTGCGAGACCCCGAGGCGACATGCTTAGACCAGTGTTCAGGCCACCGAACC 2149  
Db 2198 GGTCTCAACTGTGAGCTGGCAGGGTCCAGTCCCAGACCAAGTGTAGTGGCATGGCACT 2257  
QY 2150 TTCCTCCCGGACACCGGGCTTTGCACTGTGACCCAAAGCTGGACTGGACACGACTGTTCT 2209  
Db 2258 TACCTCCCTGACTCCGGCCCTCTGCACTGTGATCCGAACTGGATGGGTCCCAGCTGCTCT 2317  
QY 2210 ATCGAGATCTGTGCTCCGACTGTGTGGCCATGGCGTGTGCTAGGGGGCACCTGCCGC 2269  
Db 2318 GTTG--TGTGCTCAGTAGACTGTGGCACTCAGCGCTCTGCAATCGGGGGAGCCTGCCGC 2374  
QY 2270 TGCAGGATGGTGGATGGGGCAGCTGCGACCGAGCGGCTGCCACCCCGCTGTGCC 2329  
Db 2375 TGTGAAGAGGGCTGGACAGGCGCAGCTGTGACCAAGCGCTGTGCCACCCCGCTGCATT 2434  
QY 2330 GAGCATGGGACCTGCCGACGGCAAGTGCAGTGCAGCCCTGGTGGAAATGGCGAACAC 2389  
Db 2435 GAGCACGGGACCTGTAAAGATGGCAATGTGAATCCGAGAGGCTGGAATGGTGAACAC 2494  
QY 2390 TGCACCATCGTCACTATCTGATAGGGTAGTTAAAGAGGTTGCCCTGGGTGTGCAAT 2449  
Db 2495 TGCACCATTTGAT-----GGCTGCCCTGATTGTGCAAC 2527  
QY 2450 GGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTCGCTGTGCCAGCTGGCTGG 2509  
Db 2528 GGTAAACGGGAGATGCACACTGGGTCAAGAACAGCTGGCAGTGTCTGCCAGACCCGCTGG 2587  
QY 2510 AGAGGAGCTGGTGTGACACTTCCATGGAGACTGCCCTGCCGTGACAGCAAGACAATGAT 2569  
Db 2588 AGAGGCTTGGATGCAACGTTGCCATGGAAACCTCTCTGCCGTGATAACAAGGATAATGAG 2647  
QY 2570 GGAGATGGCCTGGTGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATC 2629  
Db 2648 GGAGATGGCCTGGTGGACTGCCCTGGACCCCTGACTGCTGCCCTACAGTCAAGCTGTCA 2707  
QY 2630 AACCCGCTGTGCTTGGCTCCCTAACCTCTTGACATCATCCAGGAGACACAGGTCCCT 2689  
Db 2708 AGCTGTCTGCGGGGGTCTCGGGAACCCCTTGGACATCATTCAGCAAGGTTCAGACAGAC 2767  
QY 2690 GTGTCACAGCAGAAACCTACACTCCTTCTATGACCCGATCAAGTTCCTCGTGGGCGAGGAC 2749  
Db 2768 TGGCCTGCAGTGA-----AGTCCCTTATGACCGCATCAAGCTCTTGGCAGGCAAGGAC 2821  
QY 2750 AGCAGCACATAATCCCGGGGAGAACCCCTTTTATGGAGGGGATGCTTGTGTTATTCGT 2809  
Db 2822 AGCACCACATCATCTCTGGAGACAAACCCCTTCAATAGCAGCCTGGTGTCTCTGATCCGA 2881  
QY 2810 GGCCAAGTGTACATCAGATGGAAACCCCTTGGTGGTGTGAAACATCAGTTTGTGCAAT 2869  
Db 2882 GGCCAAGTAGTAACCATGGATGGGACTCCCTTGGTGGGTGTGAAATGTGTCTTTGTCAA 2941  
QY 2870 AACCCCTCTCTTTGGATATACAAATCAGCAGGCAAGATGGCAGCTTGTGTTGACAAAT 2929  
Db 2942 TACCCAAATATGGCTACACCATCACTCGCCAGGATGGCAGCTTGTGATCCCAAT 3001  
QY 2930 GGCGGCATCTCCATCATCTGCGGTTCGAGCGGGCACCTTTTCATCACACAGGAGCACACC 2989  
Db 3002 GGGGGTCTGCTTGAATGACTTCTTCACTTTGAGCGAGCCCCCTTTTCATGAGCCAGGAGCGCACA 3061  
QY 2990 CTGTGGCTGCCATGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGGAG 3049  
Db 3062 GTGTGGCTGCCATGGAAACAGCTTCTATGCCATGGACACCCCTGTAATGAAGACCGAGGAA 3121



Db 5279 ATTGACATTGAGAACTCAACCGGGATGATGACGTCACTGTGATCACAACCTCTCCTCC 5338  
QY 5267 TCAGGCGCCTTCTACAACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCC 5326  
Db 5339 GTGGAGCCTCCTATACAGTGGTACAAGATCAAGTGCAGAACAGCTACAGCTCTGCAAT 5398  
QY 5327 GATGGCTCCTTGGGCTGCTGCTGGCCAAACGGCATGGAGTGGCGCTGCAGACTGAGCCC 5386  
Db 5399 AATGGAACCTTGGGGTGATGTACGCCAACGCCATGGCTGTGAGCTTCCACAGTGAGCCC 5458  
QY 5387 CACTTGCTGGCTGGCACCGTCAACCCCAACCGTGGCAAGAGGAATGTACAGCTGCCCATC 5446  
Db 5459 CACGTCTCGCAGGCACCATCACCCCAACCATCGGGCGTGCACATCTCTGCCCCATG 5518  
QY 5447 GACAAAGCCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGAGAGGAGTCCGGGCCAGGTC 5506  
Db 5519 GAGAAATGGCCTGAACCTCCATCGAGTGGCGCTGAGGAAGAACAGATCAAGGCAAGTC 5578  
QY 5507 ACTGTCTTTGGGCGCGGCTGCGGGTGCACAAACCGAAATCTCCTATCTCTGACTTTGAT 5566  
Db 5579 ACCATCTTTGGGAGGAGCTTCGGGTCCACGGAAGGAATCTCCTGTCCATGATTATGAC 5638  
QY 5567 CGCGTAACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTCAACCTTCGGATTCTG 5626  
Db 5639 CGAAATATCCGTACGAGAAGATCTACGATGACCAACCGGAATTCACCTTGAGGATCATC 5698  
QY 5627 TACGACAGGCGGGCGGCCAGCCTCTGTGTCACCCAGCAGCAGGCTGAATGGTGTCAAC 5686  
Db 5699 TATGACAGGTGGCGGCCCTTCTGTGGCTCCGAGCAGTGGGCTGGCAGCCGTCAAT 5758  
QY 5687 GTGACATATCTCCCTGGGGTTACATTTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGA 5746  
Db 5759 GTCTCTACTTCTCAATGGGCGCTTGGCGGCTCCAGCGAGGGCCATGAGCGAGAGG 5818  
QY 5747 ATGGAATACGACAGCGGGCGGCATCAATCCAGGATCTTCTGTCATGGGAAGACATGG 5806  
Db 5819 ACAGACATTGACAAAGCGCGGATCGTGTCCCGCATGTTTCGCCAGCGGAAAGTCTGG 5878  
QY 5807 AGCTACACATCTTAGAGAAAGTCCATGGTGTGCTACTACACAGCAGAGGCAGTATATC 5866  
Db 5879 AGTTATTCTTATCTTGACAAAGTCCATGGTCTCTGTCTACAGAGGCAACGTCAGTACATA 5938  
QY 5867 TTTGAGTTCGACAAAGATGACCGCCTCTCTTGTGACGATGCCAAACGTTGGCGGCGAG 5926  
Db 5939 TTTGAATATGACTCTCTCCGATCGCCTCCAGCAGTCACTATGCCAGTGTGCCCCGGCAC 5998  
QY 5927 AACTAGAGACCATCCGCTCAGTGGCTACTACAGAAACATCTATCAGCCCCCTGAGGSC 5986  
Db 5999 AGCATGTCCACGCACACCTCCATTGGTTACATCCGAAACATTTACAACCCACCCGAAAGC 6058  
QY 5987 AATGCTCAGTCATACAGGACTTCACTGAGGATGGGCACCTCTCTACACCTTCTACCTG 6046  
Db 6059 AATGCAATCGGTCTTTGACTACAGTACGCGCGCATCTCTAAAGACATCTTTCTTG 6118  
QY 6047 GGCAGTGGCGCAGGGTGATATACAAGTATGGCAAACTGTCAAGCTGGCAGAGACGCTC 6106  
Db 6119 GGCAGTGGCGCAGGGTGTCTACAAGTATGGAAGTCTCTCAAGTTATCAGAGATAGTC 6178  
QY 6107 TATGACACCAACCAAGGTGATTTCACTATGACGAGACGGCAGGCATGCTGAAGACCATC 6166  
Db 6179 TACGACAGCAGCCGTCACCTTTGGGTATGACGAGACCCAGGTCCTGAAGATGGTC 6238  
QY 6167 AACCTACAGAAATGAGGGCTTCACTGACCACTCCGCTACCGTCAGATTGGGCCCTGATT 6226  
Db 6239 AATCTCAAAGTGGGGCTTCTCTGTACCATCAGGTACCGAAAGGTTGGGCCCTTGTG 6298  
QY 6227 GACCGACAGATCTTCGGCTTCACTGAGGAAGGCATGGTCAACGCCCGCTTTTGACTACAAC 6286  
Db 6299 GACAAGCAGATTTACAGGTTCTCTGAGGAAGGAATGATCAACGCCAGGTTTGATTATACC 6358  
QY 6287 TAT---GACAAAGCTTCCGGGTGACAGCATGCAGGCTGTGATCAACGAGACCCCACTG 6343

Db 6359 TATCAGACAATAGCTTCCGATTGCCAGCATCAAAACCCGTCTATTAGCGAGACTCCCTT 6418  
QY 6344 CCCATTGATCTCTATCGCTATGATGTGTGAGGCAAGACAGAGCAAGTTTGGGAAGTTT 6403  
Db 6419 CCTGTTGACCTCTACCGCTATGACGAGATTTCGGCAAGGTGGAACACTTCGGCAAGTTT 6478  
QY 6404 GGTGTCAATTACTATGACATTAACAGATCATCACCAAGCTGTCTATGACCCACACCAAG 6463  
Db 6479 GGGGTCACTACTACGACATCAACAGATCATCACCACTGCCGTATGACGCTTAGCAAG 6538  
QY 6464 CATTTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTCGCTCGCTCATG 6523  
Db 6539 CACTTTGACACCCATGGGCGCATCAAGGAAGTSCAAATATGAGATGTTCCGGTCCCTCATG 6598  
QY 6524 TACTGGATGACCGTCCAGTATGATAACATGGGCGGAGTAGTGAAGAGGAGCTGAAGGTA 6583  
Db 6599 TACTGGATGACTGTGCAATATGACAGTATGGGTAGGGTTCATCAAGAGGGAACGTGAAACTA 6658  
QY 6584 GGACCCCTACGCCAATPACCACTCGCTACTCTATGAGTATGATGTCGACGGCCAGCTGCAG 6643  
Db 6659 GGGCCCTATGCCAACACACCAAGTACACCTATGACTATGACGGGACGGCCAGCTCCAG 6718  
QY 6644 ACAGTCTCCATCAATGACAAAGCCACTCTGCGCTACAGCTACGCCCTCAATGGGAACCTG 6703  
Db 6719 AGTGTGGCGCTCAATGACCGGCTTACCTGCGCTATAGCTATGACCTCAATGGGAACCTG 6778  
QY 6704 CACTTACTGAGCCCTGGAAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGAC 6763  
Db 6779 CACCTTCTAAACCCAGGAACAGTGTCTCGCTCATGCCCTTACGCTATGACCTCCGTGAC 6838  
QY 6764 CGCATCACTCGGCTGGGTGACGTGCAATACAAAGATGGATGAGGATGGCTTCTGAGGCAG 6823  
Db 6839 CGGATAACAGGCTAGGGAGCGTGCAGTACAAATCGATGACGATGGCTATTGTGCCAG 6898  
QY 6824 CGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCCTACAACCGG 6883  
Db 6899 AGAGGTCAGACATCTTTGAATACAACTCCAAGGCCCTTCTGACGAGAGCATACAACAAG 6958  
QY 6884 GCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCTGGGGCGGCGGTGTCAGGCAAG 6943  
Db 6959 GCCAGCGGATGGAGCGTGCAGTACCGCTATGACGGAGTGGGCCGCGGGCTTCTCTACAAG 7018  
QY 6944 AGCAGCCACAGCCACCACTGCGATTTCTTCTATGACAGCCTGACAAACCCCAACCAAGTC 7003  
Db 7019 ACCAACCTGGGCCACCACTACAGTACTTCTACTCCGACCTCCACAAACCCCAACAGTATC 7078  
QY 7004 ACCACCTGTACAAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGA 7063  
Db 7079 ACCCATGTTTACAACCACTCCAACCTGAGATCACCTCGCTCTACTATGACCTCCAGGGC 7138  
QY 7064 CACCTCTTTGCCATGGAGCTGAGCAGTGGTGTGATGAGTTTACATAGCTTGTGACAAACATC 7123  
Db 7139 CACCTATTTGCCATGGAGAGCAGTAGTGGTGAAGATACTATGTGCTCGCTCAGACAACACG 7198  
QY 7124 GGGACCCCTCTTGTCTGTCTTTAGTGAACAGGTTTGTATGATCAAGCAAAATCCTGTACACA 7183  
Db 7199 GGGACCCCTCTGGCTGTGTACAGTATCAATGGCCTCATGATCAAGCAAACTGCAGTACACA 7258  
QY 7184 GCCTATGGGAGATCTACATGGATACCAACCCCAACCTTTTACAGATCATATAGGCTACCAT 7243  
Db 7259 GCCTATGGGAGATCTACTATGACTCCAATCCAGACTTCCAGATGGTCTATGGCTTCCAC 7318  
QY 7244 GGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGGCGGAGATTATGATGTG 7303  
Db 7319 GGAGGCTCTATGACCCCTTACCAAGCTCGTCCACTTTTACTCAACGTGATTATGACGTG 7378  
QY 7304 CTGGCCGAGCCTGGACTAGCCCAACAGCAGGCTGTGGAAGCAGCTTAGTAGCAGCAAC 7363  
Db 7379 CTGGCAGGACGCTGGACGCTCCCCGACTACACCATGTGGAGGAACGT---GGGCAAGGAG 7435  
QY 7364 GTCATGCTTTTAACTCTCTATATGTTTCAAAAACAAACACCCCATCAGCAACTCCCAGGAC 7423  
Db 7436 CCAGCCCCCTTCAACCTGTACTGTTCAGAAACAACATCTCTGAGCAATGAGCTGGAC 7495



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:21:39 ; Search time 81 Seconds  
(without alignments)  
9650.944 Million cell updates/sec

Title: US-10-029-020-14  
Perfect score: 14887  
Sequence: 1 MDVKERKPYRSLTRRRDAER.....ELSDSANNIHFMRQSEMGR 2769

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	14887	100.0	2769	12	US-10-383-201-44
2	14887	100.0	2769	12	US-10-029-020-14
3	14610	98.1	2775	12	US-10-383-201-56
4	14529	97.6	2771	9	US-09-808-602-82
5	14529	97.6	2771	10	US-09-800-198-70
6	14406.5	96.8	2794	12	US-10-042-865-2
7	10227	68.7	2715	12	US-10-029-020-51
8	10221	68.7	2721	16	US-10-038-854-38
9	10201	68.5	2725	16	US-10-038-854-36
10	9888.5	66.4	2628	16	US-10-038-854-40
11	9856	66.2	2613	16	US-10-038-854-42
12	9616.5	64.6	2802	9	US-09-808-602-81
13	9616.5	64.6	2802	10	US-09-800-198-69
14	9616.5	64.6	2802	12	US-10-072-012-489
15	9616.5	64.6	2802	12	US-10-072-012-489

16	9610.5	64.6	2764	9	US-09-808-602-80	Sequence 80, Appl
17	9610.5	64.6	2764	10	US-09-800-198-68	Sequence 68, Appl
18	9610.5	64.6	2764	12	US-10-072-012-487	Sequence 487, App
19	9601	64.5	2765	9	US-09-808-602-84	Sequence 84, Appl
20	9601	64.5	2765	10	US-09-800-198-72	Sequence 72, Appl
21	9601	64.5	2765	12	US-10-072-012-488	Sequence 488, App
22	9584	64.4	2725	12	US-10-029-020-52	Sequence 52, Appl
23	9584	64.4	2725	16	US-10-408-765A-1687	Sequence 1687, Ap
24	9536	64.1	2733	9	US-09-808-602-8	Sequence 8, Appli
25	9536	64.1	2733	10	US-09-800-198-8	Sequence 8, Appli
26	9530.5	64.0	2724	9	US-09-808-602-13	Sequence 13, Appl
27	9530.5	64.0	2724	10	US-09-800-198-13	Sequence 13, Appl
28	9501	63.8	2725	15	US-10-295-027-928	Sequence 928, App
29	9418	63.3	2759	12	US-10-072-012-144	Sequence 144, App
30	9314	62.6	2633	15	US-10-144-194A-82	Sequence 82, Appl
31	9303	62.5	2590	12	US-10-072-012-490	Sequence 490, App
32	9145	61.4	2346	12	US-10-072-012-491	Sequence 491, App
33	8874	59.6	2551	15	US-10-144-194A-80	Sequence 80, Appl
34	6389.5	42.9	1737	9	US-09-808-602-83	Sequence 83, Appl
35	6389.5	42.9	1737	10	US-09-800-198-71	Sequence 71, Appl
36	6258.5	42.0	1688	15	US-10-144-194A-113	Sequence 113, App
37	3979.5	26.7	1045	12	US-10-042-865-54	Sequence 54, Appl
38	3979.5	26.7	1045	12	US-10-029-020-54	Sequence 54, Appl
39	3854	25.9	730	12	US-10-042-865-51	Sequence 51, Appl
40	3854	25.9	730	12	US-10-029-020-50	Sequence 50, Appl
41	3854	25.9	730	16	US-10-408-765A-2505	Sequence 2505, Ap
42	3804.5	25.6	2515	12	US-10-042-865-53	Sequence 53, Appl
43	3804.5	25.6	2515	12	US-10-029-020-53	Sequence 53, Appl
44	3730	25.1	964	12	US-10-042-865-55	Sequence 55, Appl
45	2989	20.1	768	9	US-09-773-517-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1  
US-10-383-201-44  
; Sequence 44, Application US/10383201  
; Publication No. US20040029226A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsbrook II, John et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO  
; FILE REFERENCE: 21402-568A  
; CURRENT APPLICATION NUMBER: US/10/383,201  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: 10/029020  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/365,984  
; PRIOR FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/372,022  
; PRIOR FILING DATE: 2002-04-12  
; PRIOR APPLICATION NUMBER: 60/389,143  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/391,779  
; PRIOR FILING DATE: 2002-06-26  
; PRIOR APPLICATION NUMBER: 60/410,755  
; PRIOR FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 60/412,957  
; PRIOR FILING DATE: 2002-09-23  
; PRIOR APPLICATION NUMBER: 10/051,874  
; PRIOR FILING DATE: 2002-01-16  
; PRIOR APPLICATION NUMBER: 60/366,928  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 10/055,877  
; PRIOR FILING DATE: 2002-01-22  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 44  
; LENGTH: 2769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-383-201-44

Query Match									
Best Local Similarity 100.0%; Score 14887; DB 12; Length 2769;									
Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MDVKERKPYRSLTRRDAERRYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60						
Db	1	MDVKERKPYRSLTRRDAERRYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60						
QY	61	KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDGLPQCGYSMGAGSDADMEAD	120						
Db	61	KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDGLPQCGYSMGAGSDADMEAD	120						
QY	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLRTPP	180						
Db	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLRTPP	180						
QY	181	PPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAHAQENWL	240						
Db	181	PPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAHAQENWL	240						
QY	241	LNSNIPLERNLGKQPFGLTQDNLNLIEMDILGASRHDGAYSDDHFLFKPGGTSPLFCTTS	300						
Db	241	LNSNIPLERNLGKQPFGLTQDNLNLIEMDILGASRHDGAYSDDHFLFKPGGTSPLFCTTS	300						
QY	301	PGYPLTSSVTYSPPPRPLPRSTFARPAFLNKKPKSKYCNWKCAALSAIVISATLVILLAYF	360						
Db	301	PGYPLTSSVTYSPPPRPLPRSTFARPAFLNKKPKSKYCNWKCAALSAIVISATLVILLAYF	360						
QY	361	VAMHLFGLNHLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKKGKGTTEGK	420						
Db	361	VAMHLFGLNHLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKKGKGTTEGK	420						
QY	421	PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRQVFIHPVHLKFNVSLGKAALVGIY	480						
Db	421	PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRQVFIHPVHLKFNVSLGKAALVGIY	480						
QY	481	RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSPSHETGFIQYLDSGIWH	540						
Db	481	RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSPSHETGFIQYLDSGIWH	540						
QY	541	LAFYNDGKSEVVSFLTTAIESVDNCPSPCYNGDCISGTCCHFLGFLGPDGGRASCPVL	600						
Db	541	LAFYNDGKSEVVSFLTTAIESVDNCPSPCYNGDCISGTCCHFLGFLGPDGGRASCPVL	600						
QY	601	CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTICITGTICINPGYKGSCEEV	660						
Db	601	CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTICITGTICINPGYKGSCEEV	660						
QY	661	DCMDPTCSGRGVCVRGECFCFVGWGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPSWT	720						
Db	661	DCMDPTCSGRGVCVRGECFCFVGWGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPSWT	720						
QY	721	GHDCSIEICAAACGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCTRDGKCECSPG	780						
Db	721	GHDCSIEICAAACGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCTRDGKCECSPG	780						
QY	781	WNGEHCITIAHYLDRWKKEGCPGLCNGNGRCTLDLNGWHCVQQLGWRGAGCDTSMETACGD	840						
Db	781	WNGEHCITIAHYLDRWKKEGCPGLCNGNGRCTLDLNGWHCVQQLGWRGAGCDTSMETACGD	840						
QY	841	SKDNDGDLVDCMDPDCCQLPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKF	900						
Db	841	SKDNDGDLVDCMDPDCCQLPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKF	900						
QY	901	LVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGNISFVNPLFGYTTISRDGGSF	960						
Db	901	LVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGNISFVNPLFGYTTISRDGGSF	960						
QY	961	DLVTNGGSIILRRFERAPFITQEHITLWLPWDRFFVVMETIIMRHEENEIIPSCDLSNFARP	1020						
Db	961	DLVTNGGSIILRRFERAPFITQEHITLWLPWDRFFVVMETIIMRHEENEIIPSCDLSNFARP	1020						
QY	1021	PVWSPSPLTSFASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT	1080						

Db	1021	PVWSPSPLTSFASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT	1080						
QY	1081	HPTIPFNLMMKVHLMVAVEGRLFRKWFAPAAADLSYFIWTKTDVYNQKVFLSEAFVSQV	1140						
Db	1081	HPTIPFNLMMKVHLMVAVEGRLFRKWFAPAAADLSYFIWTKTDVYNQKVFLSEAFVSQV	1140						
QY	1141	EYESCPDLILWEKRTTTLQGYEIDASKLGGWSLDKXHALNIQSGILHKGNGENQFVSQOP	1200						
Db	1141	EYESCPDLILWEKRTTTLQGYEIDASKLGGWSLDKXHALNIQSGILHKGNGENQFVSQOP	1200						
QY	1201	PVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFDFNYIRRIFFPSGNV	1260						
Db	1201	PVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFDFNYIRRIFFPSGNV	1260						
QY	1261	TNILELRNKDFRSHSHSPAHHYYLATDPMSCGAVFLSDSNRRVPFKISTVVVKDLVKNSEV	1320						
Db	1261	TNILELRNKDFRSHSHSPAHHYYLATDPMSCGAVFLSDSNRRVPFKISTVVVKDLVKNSEV	1320						
QY	1321	VAGTGDCQLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDONGIIST	1380						
Db	1321	VAGTGDCQLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDONGIIST	1380						
QY	1381	LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI	1440						
Db	1381	LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI	1440						
QY	1441	VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVTTTS	1500						
Db	1441	VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVTTTS	1500						
QY	1501	GEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRI	1560						
Db	1501	GEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRI	1560						
QY	1561	RFIRKNKPFILNTQNMVELSSPIDOELYLFDTTGKHLYTQSLPTGDYLYNFTYTGDGDTL	1620						
Db	1561	RFIRKNKPFILNTQNMVELSSPIDOELYLFDTTGKHLYTQSLPTGDYLYNFTYTGDGDTL	1620						
QY	1621	ITDNNGNMNVRRDSTGMPLWLVPDQGVYVWTMTGNSALKSVTTQGHELAMTYHGNNG	1680						
Db	1621	ITDNNGNMNVRRDSTGMPLWLVPDQGVYVWTMTGNSALKSVTTQGHELAMTYHGNNG	1680						
QY	1681	LLATKSNENGTTFYEYDSFGRLTNVTFTPTQVSSFRSDTSSVHVQVETSSKDDVTITT	1740						
Db	1681	LLATKSNENGTTFYEYDSFGRLTNVTFTPTQVSSFRSDTSSVHVQVETSSKDDVTITT	1740						
QY	1741	NLSASGAFYTLLODQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKENV	1800						
Db	1741	NLSASGAFYTLLODQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKENV	1800						
QY	1801	TLPIDNGLNLVEWRQKEQARGQVTFGRRLRVHNRNLLSLDFDRTVTRTEKIYDDHKKFT	1860						
Db	1801	TLPIDNGLNLVEWRQKEQARGQVTFGRRLRVHNRNLLSLDFDRTVTRTEKIYDDHKKFT	1860						
QY	1861	LRILYDQAGRPSLWSPSSRLNGVNVITYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFAD	1920						
Db	1861	LRILYDQAGRPSLWSPSSRLNGVNVITYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFAD	1920						
QY	1921	GKTWSYTYLEKSMVLLLSHQRYIFEFDKNDRLSSVTMPNVARQTLTIRSVGYRYNIYQ	1980						
Db	1921	GKTWSYTYLEKSMVLLLSHQRYIFEFDKNDRLSSVTMPNVARQTLTIRSVGYRYNIYQ	1980						
QY	1981	PPEGNASVIQDFTEDGHLHHTFYLGTGRRVYIKYGLSKLAETLYDTTKVSFTYDETAGM	2040						
Db	1981	PPEGNASVIQDFTEDGHLHHTFYLGTGRRVYIKYGLSKLAETLYDTTKVSFTYDETAGM	2040						
QY	2041	LKTINLQNEGFTCTIRYRQIGBLIDRQIFRFTTEEGMVNARFDYNDNSFRVTSMQAVINE	2100						
Db	2041	LKTINLQNEGFTCTIRYRQIGBLIDRQIFRFTTEEGMVNARFDYNDNSFRVTSMQAVINE	2100						
QY	2101	TPLPIDLYRYDDVSGKTEQFGKPGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYEIFR	2160						

Db 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITAVMTHTKHFDAYGRMKVEQYEIFR 2160  
QY SLMYWMTVQYDNNGRVVKELKVGYPYANTTRYSEYEDADGOLQTVSINDKPLWRYSYDLN 2220  
Db SLMYWMTVQYDNNGRVVKELKVGYPYANTTRYSEYEDADGOLQTVSINDKPLWRYSYDLN 2220  
QY GNLHLLSPGNSARLTPLRYDIRDRITRLGDVOYKXMDGDFLRQRGDIFEYNSAGLLIKA 2280  
Db GNLHLLSPGNSARLTPLRYDIRDRITRLGDVOYKXMDGDFLRQRGDIFEYNSAGLLIKA 2280  
QY YNRAGWSVRVRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLYYD 2340  
Db YNRAGWSVRVRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLYYD 2340  
QY LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAIGEYMDTNPFIQIII 2400  
Db LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAIGEYMDTNPFIQIII 2400  
QY GYHGGLYDPLTKLVHMGRDDYDLAGRWTSPDHELWKHLSSNVMPFNLYMFKNNNPISN 2460  
Db GYHGGLYDPLTKLVHMGRDDYDLAGRWTSPDHELWKHLSSNVMPFNLYMFKNNNPISN 2460  
QY SQDIKCFMTDVNSWLLTFGFLQHNVIPGYPKPDMDAMEPSYELIHTQWKTQEWDNKSIL 2520  
Db SQDIKCFMTDVNSWLLTFGFLQHNVIPGYPKPDMDAMEPSYELIHTQWKTQEWDNKSIL 2520  
QY GVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKKFASSGSVFGKGVKFKDGRVT 2580  
Db GVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKKFASSGSVFGKGVKFKDGRVT 2580  
QY TDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHTYFVKPGPSEGDLAILGLSGGRR 2640  
Db TDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHTYFVKPGPSEGDLAILGLSGGRR 2640  
QY LENGVNVTVSQINTVLNGSTRRYTDIQLQYGALCINTRYGTTLDEEKARVLELARQAVR 2700  
Db LENGVNVTVSQINTVLNGSTRRYTDIQLQYGALCINTRYGTTLDEEKARVLELARQAVR 2700  
QY QAWAREQORLREGEGELRAWTEGEKQOVLSTGRVQGYDGFVVISVEQYPELSDSANNIHF 2760  
Db QAWAREQORLREGEGELRAWTEGEKQOVLSTGRVQGYDGFVVISVEQYPELSDSANNIHF 2760  
QY MRQSEMGR 2769  
Db MRQSEMGR 2769

RESULT 2

US-10-029-020-14  
; Sequence 14, Application US/10029020  
; Publication No. US20040033971A1  
; GENERAL INFORMATION:  
; APPLICANT: Gangolli et al.  
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-225  
; CURRENT APPLICATION NUMBER: US/10/029,020  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/256,704  
; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 60/311,590  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/257,314  
; PRIOR FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: 60/311,613  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/315,617  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/307,506  
; PRIOR FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 60/322,358  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/294,075  
; PRIOR FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: 60/288,153  
; PRIOR FILING DATE: 2001-05-02  
; NUMBER OF SEQ ID NOS: 190  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 2769  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-029-020-14  
  
Query Match 100.0%; Score 14887; DB 12; Length 2769;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
Db 1 MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
QY 61 KDIVPOEAEFEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCQGYSMGAGSDADMEAD 120  
Db 61 KDIVPOEAEFEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCQGYSMGAGSDADMEAD 120  
QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLLTDEHEHENTETDHPGGLQNHARLRTPP 180  
Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLLTDEHEHENTETDHPGGLQNHARLRTPP 180  
QY 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQEPAAHAQENWL 240  
Db 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQEPAAHAQENWL 240  
QY 241 LNSNIPLETNRNLGKQPFGLTQDNLIEMDILGASHRDCAYS DGHFLFKPGGTSPLFCCTS 300  
Db 241 LNSNIPLETNRNLGKQPFGLTQDNLIEMDILGASHRDCAYS DGHFLFKPGGTSPLFCCTS 300  
QY 301 PGYPLTSSTVYSPPPRPLPRSTFARPAFNLKKPSKYCNWKCAALSAIVISATLVILLAYF 360  
Db 301 PGYPLTSSTVYSPPPRPLPRSTFARPAFNLKKPSKYCNWKCAALSAIVISATLVILLAYF 360  
QY 361 VAMHLFGLNWLQPMEGQMYEITEDTASSWPVPTDVSILYPSGGTGLETDPDRKGKGTTEGK 420  
Db 361 VAMHLFGLNWLQPMEGQMYEITEDTASSWPVPTDVSILYPSGGTGLETDPDRKGKGTTEGK 420  
QY 421 PSSFFPEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIDHPVHLKFNVSGLKAALVGIYG 480  
Db 421 PSSFFPEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIDHPVHLKFNVSGLKAALVGIYG 480  
QY 481 RKGLPPSHTQDFVELLDGRLLTQEARSLGTPRQSRGTVPSPSSHETGFIQYLDSGIWH 540  
Db 481 RKGLPPSHTQDFVELLDGRLLTQEARSLGTPRQSRGTVPSPSSHETGFIQYLDSGIWH 540  
QY 541 LAFYNDGKSEVVSFLTATAESVDNCPNSCYNGDCISGTCHCFGLGFLGPDCCGRASCPL 600  
Db 541 LAFYNDGKSEVVSFLTATAESVDNCPNSCYNGDCISGTCHCFGLGFLGPDCCGRASCPL 600  
QY 601 CSGNGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTCTITGTCICNPGYKGESCEEV 660  
Db 601 CSGNGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTCTITGTCICNPGYKGESCEEV 660  
QY 661 DCMPTCSGRGVCVRGECHCFVWGNGTNCETPRATCLDQCSGHGTFPLPDTGLCSCDPSWT 720  
Db 661 DCMPTCSGRGVCVRGECHCFVWGNGTNCETPRATCLDQCSGHGTFPLPDTGLCSCDPSWT 720  
QY 721 GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECPG 780  
Db 721 GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECPG 780  
QY 781 WNGEHTIAHYLDRVVKEGCPGLCNGNRCRTLDLNGMHCVCQLGWRGAGCDTSMETACGD 840  
Db 781 WNGEHTIAHYLDRVVKEGCPGLCNGNRCRTLDLNGMHCVCQLGWRGAGCDTSMETACGD 840  
QY 841 SKDNDGDLVDCMDPDCCCLQPLCHINPLCLGSPNPLDIIBETQVVPVSQQNLHSFYDRIKF 900  
Db 841 SKDNDGDLVDCMDPDCCCLQPLCHINPLCLGSPNPLDIIBETQVVPVSQQNLHSFYDRIKF 900

QY 901 LVGRDSTHIIIGENPFDDGGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRQDSF 960  
Db 901 LVGRDSTHIIIGENPFDDGGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRQDSF 960  
QY 961 DLVTNGGISIIILPERAPFITQEHHTLWLPWDRFFVMETIIMRHENEIPSCDLSNFARN 1020  
Db 961 DLVTNGGISIIILPERAPFITQEHHTLWLPWDRFFVMETIIMRHENEIPSCDLSNFARN 1020  
QY 1021 PVVSPSPLTSFASCAEKGPVPEIQALQBEISISGCKMRLSYLSSRTPGYKSVLRISLT 1080  
Db 1021 PVVSPSPLTSFASCAEKGPVPEIQALQBEISISGCKMRLSYLSSRTPGYKSVLRISLT 1080  
QY 1081 HPTIPFNLKMHLMVAVEGRLFRKWFAPDLSYFIWDKTDVYNQKVFLSEAFVSQV 1140  
Db 1081 HPTIPFNLKMHLMVAVEGRLFRKWFAPDLSYFIWDKTDVYNQKVFLSEAFVSQV 1140  
QY 1141 EYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHHALNIQSILHKGNGENQFVSQOP 1200  
Db 1141 EYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHHALNIQSILHKGNGENQFVSQOP 1200  
QY 1201 PVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPSGNV 1260  
Db 1201 PVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPSGNV 1260  
QY 1261 TNILELRNKDFRSHSPAHKYYLATDPMGSAVFLSDSNSRRVFKIKSTVVVKDLVKNSEV 1320  
Db 1261 TNILELRNKDFRSHSPAHKYYLATDPMGSAVFLSDSNSRRVFKIKSTVVVKDLVKNSEV 1320  
QY 1321 VAGTGQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLLYFVDGTMIRRIDQNGIIST 1380  
Db 1321 VAGTGQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLLYFVDGTMIRRIDQNGIIST 1380  
QY 1381 LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDDNNVVLQISENHQVRI 1440  
Db 1381 LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDDNNVVLQISENHQVRI 1440  
QY 1441 VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS 1500  
Db 1441 VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS 1500  
QY 1501 GEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLTNPSSLAVCADGELYVADLGNIRI 1560  
Db 1501 GEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLTNPSSLAVCADGELYVADLGNIRI 1560  
QY 1561 RFIRKNKPFINTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDYLYNFTYTGDDITL 1620  
Db 1561 RFIRKNKPFINTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDYLYNFTYTGDDITL 1620  
QY 1621 ITDNNGNMNVRRDSTGMPLWLVPDQGVYVVTMGTSNALKSVTTQGHELAMTYHGNSG 1680  
Db 1621 ITDNNGNMNVRRDSTGMPLWLVPDQGVYVVTMGTSNALKSVTTQGHELAMTYHGNSG 1680  
QY 1681 LLATKSNENGWTTFYEYDSFGRLTNVTFFPTQVSSFRSDTSSVHVQVETSSKDDVTITT 1740  
Db 1681 LLATKSNENGWTTFYEYDSFGRLTNVTFFPTQVSSFRSDTSSVHVQVETSSKDDVTITT 1740  
QY 1741 NLSASGAFYTLLODQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKENV 1800  
Db 1741 NLSASGAFYTLLODQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKENV 1800  
QY 1801 TLPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFT 1860  
Db 1801 TLPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFT 1860  
QY 1861 LRILYDQAGRPSSLWSPSSRLNGVNTYSPGGYIAGIQGIMSERMEYDQAGRITSRIFAD 1920  
Db 1861 LRILYDQAGRPSSLWSPSSRLNGVNTYSPGGYIAGIQGIMSERMEYDQAGRITSRIFAD 1920  
QY 1921 GKTWSYTYLEKSMVLLLLHSQRYIFEFDXNDRLSSVTMPNVARQTLTIRSVGYRYRNIYQ 1980  
Db 1921 GKTWSYTYLEKSMVLLLLHSQRYIFEFDXNDRLSSVTMPNVARQTLTIRSVGYRYRNIYQ 1980

QY 1981 PPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGLSKLAETLYDTTKVSFTYDETAGM 2040  
Db 1981 PPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGLSKLAETLYDTTKVSFTYDETAGM 2040  
QY 2041 LKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMNARFDYNDNSFRVTSMQAVINE 2100  
Db 2041 LKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMNARFDYNDNSFRVTSMQAVINE 2100  
QY 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYEIIFR 2160  
Db 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYEIIFR 2160  
QY 2161 SLMYWMTVQYDNMGRVVKKELKVGYPYANTTRYSEYDADGQLQTVSINDKPLWRYSYDLN 2220  
Db 2161 SLMYWMTVQYDNMGRVVKKELKVGYPYANTTRYSEYDADGQLQTVSINDKPLWRYSYDLN 2220  
QY 2221 GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFIQRGGDIFEYNSAGLLIKA 2280  
Db 2221 GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFIQRGGDIFEYNSAGLLIKA 2280  
QY 2281 YNRAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPVKVTHLYNHSSEITSLYYD 2340  
Db 2281 YNRAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPVKVTHLYNHSSEITSLYYD 2340  
QY 2341 LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKOIlyTAYGEIYMDTNPNFQIII 2400  
Db 2341 LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKOIlyTAYGEIYMDTNPNFQIII 2400  
QY 2401 GYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHKLWKLSSSNVMPFNLYMFKNNPISN 2460  
Db 2401 GYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHKLWKLSSSNVMPFNLYMFKNNPISN 2460  
QY 2461 SODIKCFMTDVNSWLLTTFGQLHNVIIPGYPKPMDMAMEPSYELIHTQMKTOEWDNSKSIL 2520  
Db 2461 SODIKCFMTDVNSWLLTTFGQLHNVIIPGYPKPMDMAMEPSYELIHTQMKTOEWDNSKSIL 2520  
QY 2521 GVQCEVQKOLKAFVTLERFDQLYGSTITSQQAPKTKKFASSGSVFGKGVKFALKDGRVT 2580  
Db 2521 GVQCEVQKOLKAFVTLERFDQLYGSTITSQQAPKTKKFASSGSVFGKGVKFALKDGRVT 2580  
QY 2581 TDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGRT 2640  
Db 2581 TDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGRT 2640  
QY 2641 LENGVNVTVSQINTVLNGRTRRTYDIOLOYGALCLNTRYGTTLDEEKARVLELARQAVR 2700  
Db 2641 LENGVNVTVSQINTVLNGRTRRTYDIOLOYGALCLNTRYGTTLDEEKARVLELARQAVR 2700  
QY 2701 QAWAREQORLREGEGLRAWTEGKQQVLSSTGRVQYDGFVVISVEQYPELSDSANNIHF 2760  
Db 2701 QAWAREQORLREGEGLRAWTEGKQQVLSSTGRVQYDGFVVISVEQYPELSDSANNIHF 2760  
QY 2761 MRQSEMGR 2769  
Db 2761 MRQSEMGR 2769

RESULT 3  
US-10-383-201-56  
; Sequence 56, Application US/10383201  
; Publication No. US20040029226A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook II, John et al.  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-568A  
; CURRENT APPLICATION NUMBER: US/10/383,201  
; CURRENT FILING DATE: 2003-03-06  
; PRIOR APPLICATION NUMBER: 10/029020  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 60/365,984  
; PRIOR FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/372,022  
; PRIOR FILING DATE: 2002-04-12

```
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 56
; LENGTH: 2775
; TYPE: PR1.
; ORGANISM: Homo sapiens
US-10-383-201-56
```

Query Match		98.1%;	Score 14610;	DB 12;	Length 2775;				
Best Local Similarity		98.2%;	Pred. No. 0;						
Matches 2743;		Conservative	2;	Mismatches 6;	Indels 42; Gaps 12;				
Qy	1	MDVKERKPYRSLTRRRDAERRYTS	SSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60					
Db	1	MDVKERKPYRSLTRRRDAERRYTS	SSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60					
Qy	61	KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDIGLPQCGYSMGAGSDADMEAD	120						
Db	61	KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDIGLPHCGYSMGAGSDADMEAD	120						
Qy	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTTEHENTETDHPGGLQNHARLRTPP	180						
Db	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTTEHENTETDHPGGLQNHARLRTPP	180						
Qy	181	PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGGAQEPHAQENWL	240						
Db	181	PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGGAQEPHAQENWL	240						
Qy	241	LNSNIPLERNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS	300						
Db	241	LNSNIPLERNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS	300						
Qy	301	PGYPLTSSTVYSPPRPLPRSTFARPAFNLKKPSKXCNWKCAALS	AIVISATLVILLAYF	360					
Db	301	PGYPLTSSTVYSPPRPLPRSTFAWPAFNLKKPSKXCNWKCAALS	AIVISATLVILLAYF	360					
Qy	361	VAMHLFGLNWHLQPMEGQMYEIT	EDTASSWPVPTDVS	LYPSGGTGLETPDRKKGKGTGK	420				
Db	361	VAMHLFGLNWHLQPMEGQMYEIT	EDTASSWPVPTDVS	LYPSGGTGLETPDRKKGKGTGK	420				
Qy	421	PSSFFPEDSFIDSGEIDVGRASQKIPPGTFWRSQVFI	DHPVHLKFNVSLGKAALVGIY	480					
Db	421	PSSFFPEASFIDSGEIDVGRASQKIPPGTFWRSQVFI	DHPVHLKFNVSLGKAALVGIY	480					
Qy	481	RKGLPPSHTQDFVELLDGRLLTQEARSL	EGTPRQSRGTVP	PPSSHETGFIQYLDSGIWH	540				
Db	481	RKGLPPSHTQDFVELLDGRLLTQEARSL	EGTPRQSRGTVP	PPSSHETGFIQYLDSGIWH	540				
Qy	541	LAFYNDGKESEVVSFLT	TAIESVDNCP	SNCYNGDCISGTCHCFLGFLGPD	CGRASC	PVL	600		
Db	541	LAFYNDGKESEVVSFLSP	-	ESVDNCP	SNCYNGDCISGTCHCFLGFLGPD	CGRASC	PVL	598	
Qy	601	CSGNGQYMKGRCLCHSGWKGAEC	CDVPTNQCID	VACSNHGT	CITGT	CICNPGYK	GESCEEV	660	
Db	599	CSGNGQYMKGRCLCHSGWKGAEC	CDVPTNQCID	VACSNHGT	CITGT	CICNPGYK	GESCEE	-	657
Qy	661	DCMDPTCSGRGVCVRGECHCFV	GWGGTNCET	PRATCLD	QCSGHG	TFLPDT	GLCSCD	PSWT	720
Db	658	DCMDPTCSGRGVCVRGECHCS	VGWGGTNCET	PRATCLD	QCSGHG	TFLPDT	GLCSCD	PSWT	717

Qy	721	GHDCSIEICAADCGGHGVCVGGT	CRCE	DGWMGAACDQ	RACHPRCAEHGT	CRD	GKCECSPG	780																																																		
Db	718	GHDCSIEICAADCGGHGVCVGGT	CRCE	DGWMGAACDQ	RACHPRCAEHGT	CRD	GKCECSPG	777																																																		
Qy	781	WNGEHCTI--AHYLD	RVVK--EGCPGL	CNGNGR	CTLDLNGWHCV	COLGWRGAGCDT	SMET	836																																																		
Db	778	WNGEHCTISLAHYL	DRVVKLS	EGCPGL	CNGNGR	CTLDLNGWHCV	COLGWRGAGCDT	SMET	837																																																	
Qy	837	ACGDSKDN	DGDLVDCMD	PDCC	LQPLCHIN	PLCLGSP	NPLDIIQETQ	VPVSQQNLHSFYD	896																																																	
Db	838	ACGDSKDN	DGDLVDCMD	PDCC	LQPLCHIN	PLCLGSP	NPLDIIQETQ	VPVSQQNLHSFYD	897																																																	
Qy	897	RIKFLVGR	SDSTHII	IPGEN	PDFG	HACVIR	QOVMTSDG	TPLVGVNISFVNNPLFGYTTISRQ	956																																																	
Db	898	RIKFLVGR	SDSTHII	IPGEN	PDFG	HACVIR	QOVMTSDG	TPLVGVNISFVNNPLFGYTTISRQ	957																																																	
Qy	957	DGSF	DLVTNGG	ISII	LRFER	APFITQ	EH	TLWLPWDRFFV	METIMRHEENE	IPSCDLSNF	1016																																															
Db	958	DGSF	DLVTNGG	ISII	LRFER	APFITQ	EH	TLWLPWDRFFV	METIMRHEENE	IPSCDLSNF	1017																																															
Qy	1017	ARPNPV	SPSPLTS	FASSCAE	KGP	IVPEIQALQ	EEISISGCK	MRLSYLSSRTPGYKSVLR	1076																																																	
Db	1018	ARPNPV	SPSPLTS	FASSCAE	KGP	IVPEIQALQ	EEISISGCK	MRLSYLSSRTPGYKSVLR	1077																																																	
Qy	1077	ISLTHPT	IPFNLMKVHLM	VAVEGR	LFRKWFAA	APDLSYFI	W	KTDVYNKQVFLSEAFV	1136																																																	
Db	1078	ISLTHPT	IPFNLMKVHLM	VAVEGR	LFRKWFAA	APDLSYFI	W	KTDVYNKQVFLSEAFV	1137																																																	
Qy	1137	SVGYE	YESCPDLIL	WEKRTTVLQGYE	IDASKLGGWSL	DKH	HALNIQS-GILHKGNGENQF	1195																																																		
Db	1138	SVGYE	YESCPDLIL	WEKRTTVLQGYE	IDASKLGGWSL	DKH	HALNIQSGGILHKGNGENQF	1197																																																		
Qy	1196	VSQPPV	IGSIMGNRRRS	ISCPSC	NGLADGNKLL	APVALT	CGSDGSLYV	GD	FN	YIR	1255																																															
Db	1198	VSQPPV	IGSIMGNRRRS	ISCPSC	NGLADGNKLL	APVALT	CGSDGSLYV	GD	FN	YIR	1257																																															
Qy	1256	PSGNVT	NI	LEL--RNK	D	FRHSHS	PAKYYLAT	D	PMSG	AVFLSD	NSRRVF	KIKSTVVVKD	1313																																													
Db	1258	PSGNVT	NI	LEL	VRNKD	FRHSHS	PAKYYLAT	D	PMSG	AVFLSD	NSRRVF	KIKSTVVVKD	1317																																													
Qy	1314	LVKNSE	V	V	A	G	T	G	D	Q	C	L	P	F	D	D	T	R	C	G	D	G	G	K	A	T	E	A	T	L	T	N	P	--GIT	V	D	K	F	L	I	Y	F	V	D	G	T	M	I	R	1370								
Db	1318	LVKNSE	V	V	A	G	T	G	D	Q	C	L	P	F	D	D	T	R	C	G	D	G	G	K	A	T	E	A	T	L	T	N	P	R	G	P	G	I	T	V	D	K	F	L	I	Y	F	V	D	G	T	M	I	R	1377			
Qy	1371	RIDQNG	I	I	S	T	L	G	S	N	D	L	S	A	R	P	L	S	C	D	S	V	M	D	I	S	O	V	R	--L	E	W	P	T	D	L	A	I	N	P	M	D	N	S	L	Y	V	L	D	N	N	1427						
Db	1378	RIDQNG	I	I	S	T	L	G	S	N	D	L	S	A	R	P	L	S	C	D	S	V	M	D	I	S	O	V	R	--L	E	W	P	T	D	L	A	I	N	P	M	D	N	S	L	Y	V	L	D	N	N	1437						
Qy	1428	VVLQISE	N	H	O	V	R	I	V	A	G	R	P	H	C	O	V	P	G	I	D	H	F	L	L	S	K	V	A	I	H	A	T	L	E	S	A	T	A	L	A	V	S	H	N	G	V	L	I	A	E	T	D	1487				
Db	1438	VVLQISE	N	H	O	V	R	I	V	A	G	R	P	H	C	O	V	P	G	I	D	H	F	L	L	S	K	V	A	I	H	A	T	L	E	S	A	T	A	L	A	V	S	H	N	G	V	L	I	A	E	T	D	1497				
Qy	1488	EKKIN	R	I	R	O	V	T	T	S	G	E	I	S	L	V	A	G	A	P	S	G	C	D	C	K	N	D	A	N	C	D	C	F	S	G	D	D	G	Y	A	K	D	A	K	L	N	T	P	S	S	L	A	V	C	A	D	1547
Db	1498	EKKIN	R	I	R	O	V	T	T	S	G	E	I	S	L	V	A	G	A	P	S	G	C	D	C	K	N	D	A	N	C	D	C	F	S	G	D	D	G	Y	A	K	D	A	K	L	N	T	P	S	S	L	A	V	C	A	D	1557
Qy	1548	GELYV	A	D	L	G	N	I	R	I	R	F	I	R	K	N	K	P	F	L	N	T	Q	N	M	Y	E	L	S	S	P	I	D	O	E	L	Y	L	F	D	T	T	G	K	H	L	Y	T	Q	S	L	P	T	G	D	Y	L	1607
Db	1558	GELYV	A	D	L	G	N	I	R	I	R	F	I	R	K	N	K	P	F	L	N	T	Q	N	M	Y	E	L	S	S	P	I	D	O	E	L	Y	L	F	D	T	T	G	K	H	L	Y	T	Q	S	L	P	T	G	D	Y	L	1617
Qy	1608	YNFTY	T	G	D	G	D	I	T	L	I	T	D	N	N	G	M	V	N	R	R	D	S	T	G	M	P	L	W	L	V	V	P	D	G	Q	V	Y	V	T	M	G	T	N	S	A	L	K	S	V	T	T	Q	G	1667			
Db	1618	YNFTY	T	G	D	G	D	I	T	L	I	T	D	N	N	G	M	V	N	R	R	D	S	T	G	M	P	L	W	L	V	V	P	D	G	Q	V	Y	V	T	M	G	T	N	S	A	L	K	S	V	T	T	Q	G	1677			
Qy	1668	HELAM	T	Y	H	G	N	S	G	L	L	A	T	K	S	N	E	N	G	W	T	T	F	Y	E	D	S	F	G	R	L	T	N	V	T	F	P	T	G	O	V	S	S	F	R	S	D	T	D	S	S	V	H	V	Q	1727		
Db	1678	HELAM	T	Y	H	G	N	S	G	L	L	A	T	K	S	N	E	N	G	W	T	T	F	Y	E	D	S	F	G	R	L	T	N	V	T	F	P	T	G	O	V	S	S	F	R	S	D	T	D	S	S	V	H	V	Q	1737		
Qy	1728	VETSS	K	D	D	V	T	I	T	N	L	S	A	S	A	F	Y	T	L	L	Q	D	O	V	R	N	S	Y	I	C	A	D	G	S	L	R	L	L	L	A	N	G	M	E	V	A	L	O	T	E	P	H	L	1787				
Db	1738	VETSS	K	D	D	V	T	I	T	N	L	S	A	S	A	F	Y	T	L	L	Q	D	O	V	R	N	S	Y	I	C	A	D	G	S	L	R	L	L	L	A	N	G	M	E	V	A	L	O	T	E	P	H	L	1793				

QY 1788 AGTNPTVGKRVNLTLPIDNGLNVEWQKEQARGQVTVFGRRLR---VHNRNLLSLDFD 1844  
DB 1794 AGTNPTVGKRVNLTLPIDNGLNVEWQKEQARGQVTVFGRRLRVLQVHNRNLLSLDFD 1853  
QY 1845 RVTRTEKIYDDHRKFTLRILYDQAGRPISLWSPSSRLNGVNVTVSPGGYIAGIQIGIMSER 1904  
DB 1854 RVTRTEKIYDDHRKFTLRILYDQAGRPISLWSPSSRLNGVNVTVSPGGYIAGIQIGIMSER 1913  
QY 1905 MEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSORQYIFEDKNDRLSSVTMPNVARQ 1964  
DB 1914 MEYDQAGRITSRIFADGKTWSYTYLEKAGV-----EFDKNDRLSSVTMPNVARQ 1962  
QY 1965 TLETIRSVGYRNIYQPPPEGNASVIQDPTEDGHLHHTFYLGTRGVYKYGKLSKLAETL 2024  
DB 1963 TLETIRSVGYRNIYQPPPEGNASVIQDPTEDGHLHHTFYLGTRGVYKYGKLSKLAETL 2022  
QY 2025 YDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGVMNAREFDYN 2084  
DB 2023 YDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGVMNAREFDYN 2082  
QY 2085 YDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAVMTHTKH 2144  
DB 2083 YDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAVMTHTKH 2142  
QY 2145 FDAYGRMKEVQEIFRSLMYMTVQYDNMGRVVKELKVGPYANTTRYSEYDADGQLQT 2204  
DB 2143 FDAYGRMKEVQEIFRSLMYMTVQYDNMGRVVKELKVGPYANTTRYSEYDADGQLQT 2202  
QY 2205 VSINDKPLWRYSDVLDNGLNHLSPGNSARLTPLRYDIRITRLGVDVQYKMDDEDGFILRQR 2264  
DB 2203 VSINDKPLWRYSDVLDNGLNHLSPGNSARLTPLRYDIRITRLGVDVQYKMDDEDGFILRQR 2262  
QY 2265 GGDIFEYNSAGLLIKAYNRAGSWSVRYRVDGLGRRVSSKSSHHLQFFYADLTNPVKVT 2324  
DB 2263 GGDIFEYNSAGLLIKAYNRAGSWSVRYRVDGLGRRVSSKSSHHLQFFYADLTNPVKVT 2322  
QY 2325 HLYNHSSEITSILYYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTA 2384  
DB 2323 HLYNHSSEITSILYYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTA 2382  
QY 2385 YGEIYMDTNPFOIIGYHGGLYDPLTKLVHMGRRDYDVLGRWTS PDHELWKHLSSNV 2444  
DB 2383 YGEIYMDTNPFOIIGYHGGLYDPLTKLVHMGRRDYDVLGRWTS PDHELWKHLSSNV 2442  
QY 2445 MPFNLYMFKNNPISNSODIKCFMTDVNSWLLTFGFQLNHVI PGYPKPDMDAMEPSYELI 2504  
DB 2443 MPFNLYMFKNNPISNSODIKCFMTDVNSWLLTFGFQLNHVI PGYPKPDMDAMEPSYELI 2502  
QY 2505 HTQMKTQEWDSK-----SILGVQCEVQKQLKAFVTLERFDQLYGSTITSCQQAOKT 2556  
DB 2503 HTQMKTQEWDSKVI PAQCQSILGVQCEVQKQLKAFVTLERFDQLYGSTITSCQQAOKT 2562  
QY 2557 KKFASSGSVFGKGVKFAKDKGRVTTDIIISVANEDGRRVAAAILNHAHYLENLHFTIDGVDI 2616  
DB 2563 KKFASSGSVFGKGVKFAKDKGRVTTDIIISVANEDGRRVAAAILNHAHYLENLHFTIDGVDI 2622  
QY 2617 HYFVKPGPSEGLAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYALCLN 2676  
DB 2623 HYFVKPGPSEGLAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYALCLN 2682  
QY 2677 TRYGTTLDEEKARVLELARQAVRQAWAREQORLREGEEGLRAWTEGEKQVVLSTGRVQG 2736  
DB 2683 TRYGTTLDEEKARVLELARQAVRQAWAREQORLREGEEGLRAWTEGEKQVVLSTGRVQG 2742  
QY 2737 YDGFVVISVEQYPELSDSANNIHFMRQSEMGRR 2769  
DB 2743 YDGFVVISVEQYPELSDSANNIHFMRQSEMGRR 2775

GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Herrman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same  
; FILE REFERENCE: 15966-697 CIP  
; CURRENT APPLICATION NUMBER: US/09/808,602  
; CURRENT FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/800,198  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 82  
; LENGTH: 2771  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-808-602-82

Query Match 97.6%; Score 14529; DB 9; Length 2771;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 2688; Conservative 40; Mismatches 41; Indels 2; Gaps 1;

QY 1 MDVKERKPYRSLTRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
DB 1 MDVKERKPYRSLTRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
QY 61 KDIVPQEAEEFCRTGANFTLRELGLLEVTTPHGTLYRTDGLPQCGYSMGAGSDADMEAD 120  
DB 61 KDMVPQEAEEFCRTGANFTLRELGLLEVTTPHGTLYRTDGLPQCGYSMGAGSDADMEAD 120  
QY 121 TVLSPEHPVPLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTTP 180  
DB 121 TVLSPEHPVPLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTTP 180  
QY 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGGAQEPAAHAQENWL 240  
DB 181 PPLPHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGGAQEPAAHAQENWL 240  
QY 241 LNSNIPLETRNLGKQPFGLTQDNLIEDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS 300  
DB 241 LNSNIPLETRNLGKQPFGLTQDNLIEDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS 300  
QY 301 PGYPLTSTVYSPPPRPLPRSTFARFAPNLKPKSKYCNWKCAALSAILISATLVILLAYF 360  
DB 301 PGYPLTSTVYSPPPRPLPRSTFARFAPNLKPKSKYCNWKCAALSAILISATLVILLAYF 360  
QY 361 VAMHLFGLNWLQPMEG--QMYEITEDTASSWPVPTDVSLYPSGGTGLTDPDRKGKGTTE 418  
DB 361 VAMHLFGLNWLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLTDPDRKGKGTTE 420  
QY 419 GKPSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFI DHPVHLKFNVS LGKAAALVGI 478  
DB 421 GKPSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFI DHPVHLKFNVS LGKAAALVGI 480  
QY 479 YGRKGLPPSHTQDFVELLDGRRLLTQEARSLLEGTPRQSRGTVPPSSSHETGFIQYLDGI 538  
DB 481 YGRKGLPPSHTQDFVELLDGRRLLTQEARSLLEGTPRQSRGTVPPSSSHETGFIQYLDGI 540  
QY 539 WHLAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCISGTCFLGLPDCGRASCP 598  
DB 541 WHLAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCISGTCFLGLPDCGRASCP 600  
QY 599 VLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTICNPGYKGESCE 658  
DB 601 VLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTICNPGYKGESCE 660

QY 659 EVDCMDPTCSGRGVCVRGECHCFVGVGGTNCETPRATCLDQCSGHHGFLPDTGLCSCDPS 718  
Db 661 EVDCMDPTCSSRGVCVRGECHCSVGVGGTNCETPRATCLDQCSGHHGFLPDTGLCNC DPS 720  
QY 719 WTGHDCSIEICAAOCGGHGVCGGTCTRCEDGWMGAACDQACHPRCAEHGTCRDGKCECS 778  
Db 721 WTGHDCSIEICAAOCGGHGVCGGTCTRCEDGWMGAACDQACHPRCAEHGTCRDGKCECS 780  
QY 779 PGWNGEHCTIAHYLDRVVVKEGCPGLCNGNGRCTLDLNGWHVCVQLGWRGAGCDTSMETAC 838  
Db 781 PGWNGEHCTIAHYLDRVVVKEGCPGLCNGNGRCTLDLNGWHVCVQLGWRGTGCDTSMETGC 840  
QY 839 GDSKNDGDGLVDCMDPCCLQPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSHFYDRI 898  
Db 841 GDGKNDGDGLVDCMDPCCLQPLCHVNPLCLGSPDPLDIIQETQAPVSQQNLNPFYDRI 900  
QY 899 KFLVGRDSTHIIIPGENPFDGGHACVIRQVMTSDGTPLVGVNISFVNNPLFGYTISRQDG 958  
Db 901 KFLVGRDSTHSIPGENPFDGGHACVIRQVMTSDGTPLVGVNISFINNPLFGYTISRQDG 960  
QY 959 SFDLVTNGGSIILRPERAPFITQEHTLWLPWDRFFVMETIIMRHEENEIPSCDLSNFAR 1018  
Db 961 SFDLVTNGGSIILRPERAPFITQEHTLWLPWDRFFVMETIIVMRHEENEIPSCDLSNFAR 1020  
QY 1019 PNPVVSPPLTSPASSCAEKGPVPEIQALQEBEISISGCKMRLSYLSSRTPGYKSVLRI 1078  
Db 1021 PNPVVSPPLTSPASSCAEKGPVPEIQALQEBEIVIAGCKMRLSYLSSRTPGYKSVLRI 1080  
QY 1079 LTHPTIPFNLKMKVHLMVAVEGRFLFRKWFAAAPDLSYFIWDKTDVYNQKVFGLSEAFVS 1138  
Db 1081 LTHPTIPFNLKMKVHLMVAVEGRFLFRKWFAAAPDLSYFIWDKTDVYNQKVFGSEAFVS 1140  
QY 1139 GYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHALNIQSGILHKNGENQFVSQ 1198  
Db 1141 GYEYESCPDLILWEKRTAVLQGYEIDASKLGWSLDKHALNIQSGILHKNGENQFVSQ 1200  
QY 1199 QPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGD FNYIRIFPSG 1258  
Db 1201 QPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGD FNYIRIFPSG 1260  
QY 1259 NVTNILELRNKDFRSHSPAHKYYLATDPMGAVFLSDSNSRRVFKTKSTVAVKDLVKNS 1318  
Db 1261 NVTNILEMRNKDFRSHSPAHKYYLATDPMGAVFLSDTNSRRVFKVKS TT VAVKDLVKNS 1320  
QY 1319 EVVAGTGDCQLPFDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDONGII 1378  
Db 1321 EVVAGTGDCQLPFDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRV DONGII 1380  
QY 1379 STLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVVLQISENHQV 1438  
Db 1381 STLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLDNNVVLQISENHQV 1440  
QY 1439 RIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVT 1498  
Db 1441 RIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVT 1500  
QY 1499 TSGEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLNTPSSLAVCADGELYVADLGN 1558  
Db 1501 TSGEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLNTPSSLAVCADGELYVADLGN 1560  
QY 1559 RIRFIRKNKPFNTQNMVELSSPIDQELYLEDTTGKHLYTQSLPTGDYLYNFYTTGDGDI 1618  
Db 1561 RIRFIRKNKPFNTQNMVELSSPIDQELYLEDTSKGHLYTQSLPTGDYLYNFYTTGDGDI 1620  
QY 1619 TLITDNNGNMNVNRDSTGMPLWLVPDGVVYVMTGMTNSALKSVTTQGHELAMTYHGN 1678  
Db 1621 THITDNNGNMNVNRDSTGMPLWLVPDGVVYVMTGMTNSALRSVTTQGHELAMTYHGN 1680  
QY 1679 SGLLATKSNENGWTTFFEYDSFGRLTNVTFPTGQVSSFRSDTSSVHVQVETSSKDDVTI 1738  
Db 1681 SGLLATKSNENGWTTFFEYDSFGRLTNVTFPTGQVSSFRSDTSSVHVQVETSSKDDVTI 1740  
QY 1739 TTNLSASGAFYTL LQDQVRNSYYIGADGSLRLLLLANGMEVALQTEPHLLAGTVNPTVGKR 1798

Db 1741 TTNLSASGAFYTL LQDQVRNSYYIGADGSLRLLLLANGMEVALQTEPHLLAGTVNPTVGKR 1800  
QY 1799 NVTLPIDNGLNLVEWRQORKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRK 1858  
Db 1801 NVTLPIDNGLNLVEWRQORKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRK 1860  
QY 1859 FTLRILYDQAGRPSLWSPSSRLNGVNVITYSPGGYIAGIQRIMSERMEYDQAGRITSRIF 1918  
Db 1861 FTLRILYDQAGRPSLWSPSSRLNGVNVITYSPGGHIAIQRIMSERMEYDQAGRITSRIF 1920  
QY 1919 ADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDLSSVTMPNVARQTLETIRSVGYRYNI 1978  
Db 1921 ADGKMSYTYTYLEKSMVLLHLSQRQYIFEFDKNDLSSVTMPNVARQTLETIRSVGYRYNI 1980  
QY 1979 YQPPEGNASVIOQFTEDEGHLLHTFYLGTGRRVYKYGKLSKLAETLYDTTKVSFTYDETA 2038  
Db 1981 YQPPEGNASVIOQFTEDEGHLLHTFYLGTGRRVYKYGKLSKLAETLYDTTKVSFTYDETA 2040  
QY 2039 GMLKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEEGMVNARFDYNDNSFRVTSMQAVI 2098  
Db 2041 GMLKTVNLQNEGFTCTIRYQIGPLIDRQIFRFTTEEGMVNARFDYNDNSFRVTSMQAVI 2100  
QY 2099 NETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYEI 2158  
Db 2101 NETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYEI 2160  
QY 2159 FRSLMYWMTVQYDNMGVVVKELKVGPYANTTRYSEYEDADGQLQTVSINDKPLWRYSYD 2218  
Db 2161 FRSLMYWMTVQYDNMGVVVKELKVGPYANTTRYSEYEDADGQLQTVSINDKPLWRYSYD 2220  
QY 2219 LINGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFLRQGGDIFEYNSAGLLI 2278  
Db 2221 LINGNLHLLSPGNSARLTPLRYDLDRITRLGDVQYKMDDEGFLRQGGDVFEYNSAGLLI 2280  
QY 2279 KAYNRAGWSVRYRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLY 2338  
Db 2281 KAYNRAGWSVRYRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLY 2340  
QY 2339 YDLQGHLFAMELSSGDEFIYACDNIGTPLAVFSGTGLMIKQILYTAIGEYINDTNPNFQI 2398  
Db 2341 YDLQGHLFAMELSSGDEFIYACDNIGTPLAVFSGTGLMIKQILYTAIGEYINDTNPNFQI 2400  
QY 2399 IIGYHGGLYDPLTKLVHMGRRDDYDLAGRWTSPDHELWKHLSSSNVMPFNLVFKNNNPI 2458  
Db 2401 IIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKRLSSNSIVPFHLYNFKNNNPI 2460  
QY 2459 SNSQDIKCFMTDVNSWLLTFGFQLHNVI PGYPKPDMDAMEPSYELIHTQMKTOEWDNSKS 2518  
Db 2461 SNSQDIKCFMTDVNSWLLTFGFQLHNVI PGYPKPDTDAMEPSYELVHTQMKTOEWDNSKS 2520  
QY 2519 ILGVQCEVQQLKAFVTLERFDQLYGSTITSQOAPKTKKFASSGVSFGKGVKFPALKDGR 2578  
Db 2521 ILGVQCEVQQLKAFVTLERFDQLYGSTITSQOAPETKKFASSGSI FGKGVKFPALKDGR 2580  
QY 2579 VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGD LAILGLSGGR 2638  
Db 2581 VTTDIISVANEDGRRIAAILNNAHYLENLHFTIDGVDTHYFVKPGPSEGD LAILGLSGGR 2640  
QY 2639 RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGA LCLNTRYGTTLDDEEKARVLELARORA 2698  
Db 2641 RTLENGVNVTVSQINTMLSGRTRRYTDIQLQYRA LCLNTRYGTTVDEEKVRVLELARORA 2700  
QY 2699 VRQAWAREQOORLREGEGELRAWTEGEKQOVLSTGRVQYDGFVVISVEQYPELSDSANNI 2758  
Db 2701 VRQAWAREQOORLREGEGELRAWTDGEKQOVLNTRGVQYDGFVFTSVEQYPELSDSANNI 2760  
QY 2759 HMRQSEMGRR 2769  
Db 2761 HMRQSEMGRR 2771

RESULT 5

US-09-800-198-70									
; Sequence 70, Application US/09800198									
; Publication No. US20030087816A1									
; GENERAL INFORMATION:									
; APPLICANT: Vernet, Cornie AM									
; APPLICANT: Fernandes, Elma									
; APPLICANT: Shimkets, Richard A									
; APPLICANT: Herrmann, John L									
; APPLICANT: Majumder, Kumud									
; APPLICANT: Mishra, Vishna									
; APPLICANT: Mezes, Peter S									
; APPLICANT: Rastelli, Luca									
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME									
; FILE REFERENCE: 15966-697									
; CURRENT APPLICATION NUMBER: US/09/800,198									
; CURRENT FILING DATE: 2001-03-05									
; PRIOR APPLICATION NUMBER: 60/186,596									
; PRIOR FILING DATE: 2000-03-03									
; NUMBER OF SEQ ID NOS: 98									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 70									
; LENGTH: 2771									
; TYPE: PRT									
; ORGANISM: Mus musculus									
US-09-800-198-70									
Query Match									
Best Local Similarity 97.6%; Score 14529; DB 10; Length 2771;									
Matches 2688; Conservative 40; Mismatches 41; Indels 2; Gaps 1;									
QY	1	MDVKERKPYRSLTRRRDAERRY	TSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60					
Db	1	MDVKERKPYRSLTRRRDAERRY	TSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60					
QY	61	KDIVPQEAEEFCRTGANFTLRELGL	EEVTPPHGTLTYRTDIGLPQCGYSMGAGSDADMEAD	120					
Db	61	KDMVPQEAEEFCRTGNTFTLRELGL	EMTPPHGTLTYRTDIGLPCHGYSMGASSADLEAD	120					
QY	121	TVLSPEHPVRLWGRSTR	SGRSSCLSSRANSNLTLTDTEHENTETDHPGGLQNHARLTPP	180					
Db	121	TVLSPEHPVRLWGRSTR	SGRSSCLSSRANSNLTLTDTEHENTETDHPSSLQNHPRLTPP	180					
QY	181	PPLSHAHTPNQHHAAASINSLNRGNFT	PRSNPSPAPTDHLSLGEPPAGGAQEPAAHAQENWL	240					
Db	181	PPLPHAHTPNQHHAAASINSLNRGNFT	PRSNPSPAPTDHLSLGEPPAGSAQEPTHAQDNWL	240					
QY	241	LNSNIPLSTRNLGKQPF	LGTQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS	300					
Db	241	LNSNIPLSTRNLGKQPF	LGTQDNLIEMDILSASRHDGAYSDGHFLFKPGGTSPLFCTTS	300					
QY	301	PGYPLTSTVYSPPPRPLPRST	FARPAFNLKPKSKYCNWKAALSAIVISATLVILLAYF	360					
Db	301	PGYPLTSTVYSPPPRPLPRST	FSPRPAFNLKPKSKYCNWKAALSAILISATLVILLAYF	360					
QY	361	VAMHFLGNLWHLQPMEG--	QMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGGTTE	418					
Db	361	VAMHFLGNLWHLQPMEGQM	QMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGGAAE	420					
QY	419	GKPSFFPDSFIDS	GEIDVGRASQKIPPGTFWRSQVFDHPVHLKFNVSGLGKAALVGI	478					
Db	421	GKPSLFPDSFIDS	GEIDVGRASQKIPPGTFWRSQVFDHPVHLKFNVSGLGKAALVGI	480					
QY	479	YGRKGLPPSHTQDF	VELLDGRRLLTQEARSLGTPRQSRGTVPPSSHETGFIQYLDSCI	538					
Db	481	YGRKGLPPSHTQDF	VELLDGRRLLTQEARSLGTPRQSRGVPVPPSSHETGFIQYLDSCI	540					
QY	539	WHLAFYNDGKESEV	SFLTATAIESVDNCPSNCYGNGDCISGTCHCFGLGPDGGRASCP	598					
Db	541	WHLAFYNDGKESEV	SFLTATAIESVDNCPSNCYGNGDCISGTCHCFGLGPDGGRASCP	600					
QY	599	VLCSGNGQYMKGRCL	CHSGWKGAFCDVPTNQCIDVACSNHGTCTITGTCICNPGYKGESCE	658					
Db	601	VLCSGNGQYMKGRCL	CHSGWKGAFCDVPTNQCIDVACSSHGTCTIMGTCTICNPGYKGESCE	660					
QY	659	EVDCMDPTCSGRGVCVR	GECHCFVGVGGTNCETPRATCLDQC	SGHGTFLPDTGLCSCDPS	718				
Db	661	EVDCMDPTCSSRGVCVR	GECHCHSVGGTNCETPRATCLDQC	SGHGTFLPDTGLCNC DPS	720				
QY	719	WTGHDCSIEICAADCG	HGVCVGGTCRCEDGWMGAACQ	RACHPRCAEHGTCRDGKCECS	778				
Db	721	WTGHDCSIEICAADCG	HGVCVGGTCRCEDGWMGAACQ	RACHPRCAEHGTCRDGKCECS	780				
QY	779	PGWNGEHTIAHYLDR	VVKEGCPGLCNGNGRCTLDLNGW	HVCVQLGWRGAGCDTSMETAC	838				
Db	781	PGWNGEHTIAHYLDR	VVKEGCPGLCNGNGRCTLDLNGW	HVCVQLGWRGTGCDTSMETGC	840				
QY	839	GDSKNDGDGLVDCMD	PDCCLOPLCHINPLCLGSPNPLDII	QETQVPVSQQNLHSPFYDRI	898				
Db	841	GDGKNDGDGLVDCMD	PDCCLOPLCHVNPCLGSPDPLDII	QETQAPVSQQNLNPFYDRI	900				
QY	899	KFLVGRDSTHII	PGENPFDDGGHACVIRGQVMTSDGT	PLVGVNISFVNNPLFGYTISRQDG	958				
Db	901	KFLVGRDSTHSI	PGENPFDDGGHACVIRGQVMTSDGT	PLVGVNISFINNPLFGYTISRQDG	960				
QY	959	SFDLVINGGISI	ILRFERAPFITQEH	TLWLPWDRFFVMEIIMRHEENEIP	SCDLSNFAR	1018			
Db	961	SFDLVINGGISI	ILRFERAPFITQEH	TLWLPWDRFFVMEIIVMRHEENEIP	SCDLSNFAR	1020			
QY	1019	PNPVVSPSLT	SFASSCAEKGIVPEIQALQEEISIS	CKMRLSYLSRSTPGYKSVLRIS	1078				
Db	1021	PNPVVSPSLT	SFASSCAEKGIVPEIQALQEEI	VIAGCMRLSYLSRSTPGYKSVLRIS	1080				
QY	1079	LTHPTIPFNL	MKVHLMVAVEGR	LFRKWFAPDLSYFIW	KTDVYNQKVFGLSEAFVS	1138			
Db	1081	LTHPTIPFNL	MKVHLMVAVEGR	LFRKWFAPDLSYFIW	KTDVYNQKVFGESEAFVS	1140			
QY	1139	GYEYESCPDL	LILWEKRTTVLQGEIDASK	LGGWSLDKHHALNIQSGIL	HKNGENQFVSQ	1198			
Db	1141	GYEYESCPDL	LILWEKRTAVLQGEIDASK	LGGWSLDKHHALNIQSGIL	HKNGENQFVSQ	1200			
QY	1199	QPPVIGSIMGNRRRS	ISCPSCNGLADGNKLLAPVAL	TCGSDGSLYVGFNYIRRI	PPSG	1258			
Db	1201	QPPVIGSIMGNRRRS	ISCPSCNGLADGNKLLAPVAL	TCGSDGSLYVGFNYIRRI	PPSG	1260			
QY	1259	NVTNILELRNKD	FRHSHSPAHHYLLATDPM	SGAVFLSDNSRRVFKIKSTVV	VVKDLVKNS	1318			
Db	1261	NVTNILEMRNKD	FRHSHSPAHHYLLATDPM	SGAVFLSDNSRRVFKVKS	TTVVVKDLVKNS	1320			
QY	1319	EVVAGTDQCL	PFDDTRCGDGGKATEATLT	NPRGITVDKFLIYFVDGT	MIRRIDQNGII	1378			
Db	1321	EVVAGTDQCL	PFDDTRCGDGGKATEATLT	NPRGITVDKFLIYFVDGT	MIRRIDQNGII	1380			
QY	1379	STLLGSNDLTSAR	PLSCDSVMDISQVRLEWPTDL	AINPMDNSLYVLDNNVVLQISE	NHQV	1438			
Db	1381	STLLGSNDLTSAR	PLSCDSVMEISQVRLEWPTDL	AINPMDNSLYVLDNNVVLQISE	NHQV	1440			
QY	1439	RIVAGRPMHCQ	VPIDHFLLSKVAIHATLES	ATALAVSHNGVLYIAETDEK	KINRIRQVT	1498			
Db	1441	RIVAGRPMHCQ	VPIDHFLLSKVAIHATLES	ATALAVSHNGVLYIAETDEK	KINRIRQVT	1500			
QY	1499	TSGEISLVAGAP	SGCDCKNDANCDCFSGDD	GYAKDAXLNTPSSLAV	CADGELYVADL	1558			
Db	1501	TSGEISLVAGAP	SGCDCKNDANCDCFSGDD	GYAKDAXLNTPSSLAV	CADGELYVADL	1560			
QY	1559	RIRFIRKNKPF	LNTQNMVELSSPIDQEL	YLFDDTTGKHL	YTSQSLPTGDLYNFT	1618			
Db	1561	RIRFIRKNKPF	LNTQNMVELSSPIDQEL	YLFDDTTGKHL	YTSQSLPTGDLYNFT	1620			
QY	1619	TLITDNNGMV	NVRRDSTGMPLWL	VVPDQVYVVTMTG	NSALKSVTTQGH	ELAMTYHGN	1678		
Db	1621	THITDNNGMV	NVRRDSTGMPLWL	VVPDQVYVVTMTG	NSALRSVTTQGH	ELAMTYHGN	1680		
QY	1679	SGLLATKSNENG	WTTTFYEYDSFGRL	NTNVFTPTGQVSSFR	SDTSSVHVQVET	SSKDDVTI	1738		
Db	1681	SGLLATKSNENG	WTTTFYEYDSFGRL	NTNVFTPTGQVSSFR	SDTSSVHVQVET	SSKDDVTI	1740		

QY	1739	TTNLSAGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLAGTVNPTVGKR	1798
Db	1741	TTTTLSAGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLAGTVNPTVGKR	1800
QY	1799	NVTLPIDNGLNLVWVRQKEQARGQVTVFGRRLRVHNRNLLSLDFRTRTEKIYDDHRK	1858
Db	1801	NVTLPIDNGLNLVWVRQKEQARGQVTVFGRRLRVHNRNLLSLDFRTRTEKIYDDHRK	1860
QY	1859	FTLRILYDQAGRPSSLWSPSSRLNGVNVTTSPGVIAGIQRGIMSERMEYDQAGRITSRIF	1918
Db	1861	FTLRILYDQAGRPSSLWSPSSRLNGVNVTTSPGVIAGIQRGIMSERMEYDQAGRITSRIF	1920
QY	1919	ADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLLSSVTMPNVARQIETIRSVGYRNI	1978
Db	1921	ADGKMWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLLSSVTMPNVARQIETIRSVGYRNI	1980
QY	1979	YQPPEGNASVIOQFTEDGHLHLLTFYLGTRRRVIYKYGKLSKLAETLYDTTKVSFTYDETA	2038
Db	1981	YQPPEGNASVIOQFTEDGHLHLLTFYLGTRRRVIYKYGKLSKLAETLYDTTKVSFTYDETA	2040
QY	2039	GMLKTNLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNARFDYNDNSFRVTSMQAVI	2098
Db	2041	GMLKTNLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNARFDYNDNSFRVTSMQAVI	2100
QY	2099	NETPLPIDLYRDDVSGKTEQFGKFGVIYYDINQIITAVMTHTKHFDAYGRMKEVQYEI	2158
Db	2101	NETPLPIDLYRDDVSGKTEQFGKFGVIYYDINQIITAVMTHTKHFDAYGRMKEVQYEI	2160
QY	2159	FRSLMYWMTVQYDNMGVVVKELKVGYPYANTTRYSEYDADGQLQTVSINDKPLWRYSYD	2218
Db	2161	FRSLMYWMTVQYDNMGVVVKELKVGYPYANTTRYSEYDADGQLQTVSINDKPLWRYSYD	2220
QY	2219	LNGNLHLLSPGNSARLTPLRYDIDRITRLGDVQYKMDDEDGFLRQGGDIFEYNSAGLLI	2278
Db	2221	LNGNLHLLSPGNSARLTPLRYDIDRITRLGDVQYKMDDEDGFLRQGGDIFEYNSAGLLI	2280
QY	2279	KAYNRAGWSVRVRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLSY	2338
Db	2281	KAYNRAGWSVRVRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLSY	2340
QY	2339	YDLQCHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYPEIYMDTNPFI	2398
Db	2341	YDLQCHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYPEIYMDTNPFI	2400
QY	2399	IIGYHGGLYDPLTKLVHMGRRDDYDLAGRWTSPDHKLWHLSSNVMPFNLYMFKNNPI	2458
Db	2401	IIGYHGGLYDPLTKLVHMGRRDDYDLAGRWTSPDHKLWHLSSNVMPFNLYMFKNNPI	2460
QY	2459	SNSQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDSKS	2518
Db	2461	SNSQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDSKS	2520
QY	2519	ILGVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKKFASSGVSFGKGVKFALKDGR	2578
Db	2521	ILGVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKKFASSGVSFGKGVKFALKDGR	2580
QY	2579	VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGR	2638
Db	2581	VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGR	2640
QY	2639	RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYCALCLNTRYGTTLDEEKARVLELARQRA	2698
Db	2641	RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYCALCLNTRYGTTLDEEKARVLELARQRA	2700
QY	2699	VRQAWAREQORLREGEEGELRAWTEGEKQOQVLSGTGRVQYDGFVIVSVEQYPELSDSANNI	2758
Db	2701	VRQAWAREQORLREGEEGELRAWTEGEKQOQVLSGTGRVQYDGFVIVSVEQYPELSDSANNI	2760
QY	2759	HFMROSEMGR 2769	
Db	2761	HFMROSEMGR 2771	

RESULT 6  
US-10-042-865-2  
; Sequence 2, Application US/10042865  
; Publication No. US20040029216A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Vernet, Corine A.M  
; APPLICANT: Taylor, Sarah  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Miller, Charles E  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Boldog, Ference L  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Rothenberg, Mark E  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Gunther, Erik  
; APPLICANT: Stone, David  
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of  
; TITLE OF INVENTION: Using the Same  
; FILE REFERENCE: 21402-537  
; CURRENT APPLICATION NUMBER: US/10/042,865  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/260,417  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/260,831  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: 60/272,338  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/274,876  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/284,704  
; NUMBER OF SEQ ID NOS: 264  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2794  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-042-865-2

Query Match 96.8%; Score 14406.5; DB 12; Length 2794;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 2704; Conservative 21; Mismatches 41; Indels 31; Gaps 11;

QY	1	MDVKERKPYRSLTRRRDAERYTSSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60
Db	1	MDVKERKPYRSLTRRRDAERYTSSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60
QY	61	KDIVPQEAEEFCRTGANFTLRELGLLEEVTPPHGTLVYRTDGLPQCGYMGAGSDADMEAD	120
Db	61	KDIVPQEAEEFCRTGANFTLRELGLLEEVTPPHGTLVYRTDGLPQCGYMGAGSDADMEAD	120
QY	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTITDTEHENTETDHPGGLQNHARLRTPP	180
Db	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTITDTEHENTETDHPGGLQNHARLRTPP	180

QY	181	PPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQPAHAQENWL	240
Db	181	PPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQPAHAQENWL	240
QY	241	LNSNIPLETRNLGQPFGLTQDNLIEMDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS	300
Db	241	LNSNIPLETRNLGQPFGLTQDNLIEMDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS	300
QY	301	PGYPLTSSTVSPRPRLPRSTFAWPAENLKKPSKYCNWKCAALSAIVISATLVILLAYF	360
Db	301	PGYPLTSSTVSPRPRLPRSTFAWPAENLKKPSKYCNWKCAALSAIVISATLVILLAYF	360
QY	361	VAMHLFGLNHLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKKGTTTEGK	420
Db	361	VAMHLFGLNHLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKKGTTTEGK	420
QY	421	PSSFFPEDSFIDSGEIDVGRRASOKIPPGTFWRSQVFIDHPVHLKFNVSLGKAALVGIYG	480
Db	421	PSSFFPEASFIDSGEIDVGRRASOKIPPGTFWRSQVFIDHPVHLKFNVSLGKAALVGIYG	480
QY	481	RKGLPPSHTQDPFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDGSIWH	540
Db	481	RKGLPPSHTQDPFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDGSIWH	540
QY	541	LAFYNDGKESVVSFLTTAIESVDNCPNSCYNGDCISGTHCFLGFLGPDGCRASCPVL	600
Db	541	LAFYNDGKESVVSFLTTAI--LDSW-ALCLGDGECVSGTCHCFPGFLGPDGCRASCPVL	597
QY	601	CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTICNPGYKGESCEEV	660
Db	598	CSGNGQYSKGRCLCFSGWKGTEDVPTTQCIDPQCGGRGICMGSCACNSGYKGESCEEA	657
QY	661	DCMDPTCSGRGVCVRGECHCFVGGWGTNCETPRATCLDQCSHGHTFLPDTGLCSDPSSWT	720
Db	658	DCIDPGCSNHGVCIHGECHCSPGWGGSNCEILKTMCPDQCSHGHTYQESGSCCTCDPNWT	717
QY	721	GHDCSIEICADCGGHGVCVGTGCTRCBEGWMAACDQACHPRCAEHGTCRDGKCECSPG	780
Db	718	GPDCSNEICSDGSHGVCMGGTCRCBEGWGPACNQACHPRCAEHGTCRDGKCECSPG	777
QY	781	WNGEHCTI--AHYLDREVVK--EGCPGLCNGNGRCTLDLNGWHCVCOLGWRGAGCDTSMET	836
Db	778	WNGEHCTISLAHYLDREVVKLEGGCPGLCNGNGRCTLDLNGWHCVCOLGWRGAGCDTSMET	837
QY	837	ACGDSKONDGDLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSPFYD	896
Db	838	ACGDSKONDGDLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSPFYD	897
QY	897	RIKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFVNPNPLFGYTIISRQ	956
Db	898	RIKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFVNPNPLFGYTIISRQ	957
QY	957	DGSFDLVTNGGSIILRFERAPFITQHTLWLPWDRFFVFMETIIMRHEENEIIPSCDLSNF	1016
Db	958	DGSFDLVTNGGSIILRFERAPFITQHTLWLPWDRFFVFMETIIMRHEENEIIPSCDLSNF	1017
QY	1017	ARNPNVVSPLTSPASSCAEKGPIVPEIQALQEEIISISGCKMRLSYLSSRTPGYKSVL	1076
Db	1018	ARNPNVVSPLTSPASSCAEKGPIVPEIQALQEEIISISGCKMRLSYLSSRTPGYKSVL	1077
QY	1077	ISLTHPTIPNLMKVHLMVAVEGRLEKFWFAAADLSYFIWDKTDVYNQKVFLSEAFV	1136
Db	1078	ISLTHPTIPNLMKVHLMVAVEGRLEKFWFAAADLSYFIWDKTDVYNQKVFLSEAFV	1137
QY	1137	SVGYEYESCPDILWEKRTTVLQGYEIDASKLGGWSLDKHALNIQS--GILHKNGENQF	1195
Db	1138	SVGYEYESCPDILWEKRTTVLQGYEIDASKLGGWSLDKHALNIQS--GILHKNGENQF	1197
QY	1196	VSQPPVIGSMNGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRF	1255
Db	1198	VSQPPVIGSMNGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRF	1257
QY	1256	PSGNVTNILEL--RNKDFRHSHPAHKYIYLATDPMMSGAVFLSDSNSRRVFKIKSTVVVKD	1313

Db	1258	PSGNVTNILELVRNKDFRHSHPAHKYIYLATDPMMSGAVFLSDSNSRRVFKIKSTVVVKD	1317
QY	1314	LVKNSEVVAGTGDOCLPFDDTRCGDGGKATEATLTNPR--GITVDKFGLIYFVDGTMIR	1370
Db	1318	LVKNSEVVAGTGDOCLPFDDTRCGDGGKATEATLTNPRGPPGIIITVDKFGLIYFVDGTMIR	1377
QY	1371	RIDQNGIISTLLGSNDLTSARPLSCDSVMDISOVR--LEWPTDLAINPMDNSLYVLDDN	1427
Db	1378	RIDQNGIISTLLGSNDLTSARPLSCDSVMDISOVRQVHLEWPTDLAINPMDNSLYVLDDN	1437
QY	1428	VVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLZIAETD	1487
Db	1438	VVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLZIAETD	1497
QY	1488	EKKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLTNPSSSLAVCAD	1547
Db	1498	EKKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLTNPSSSLAVCAD	1557
QY	1548	GELYVADLGNIRIRFIRKKNPFLNTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDYL	1607
Db	1558	GELYVADLGNIRIRFIRKKNPFLNTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDYL	1617
QY	1608	YNFTYTGDDITLITDNNGNMNVRRDSTGMPLWLVPDQGVYVVTMTGNSALKSVTTQ	1667
Db	1618	YNFTYTGDDITLITDNNGNMNVRRDSTGMPLWLVPDQGVYVVTMTGNSALKSVTTQ	1677
QY	1668	HELAMMTYHGNSGLLATKSNENGTTFEYDVSFGRLTNVFTPTQVSSFSRSDTSSSVHVQ	1727
Db	1678	HELAMMTYHGNSGLLATKSNENGTTFEYDVSFGRLTNVFTPTQVSSFSRSDTSSSVHVQ	1737
QY	1728	VETSSKDDVTITTNLSASGAFYTLQDQVRNSYIIGADGSLRLLLANGMEVALQTEPHLL	1787
Db	1738	VETSSKDDVTITTNLSASGAFYTLQDQVRNSYIIGADGSLRLLLANGMEVALQTEPHLL	1797
QY	1788	AGTVNPTVGRNVTLPIDNGLNLVWRQKEQARGQVTVFGRRLR--VHNRNLLSLDDFD	1844
Db	1798	AGTVNPTVGRNVTLPIDNGLNLVWRQKEQARGQVTVFGRRLRVLQVHNRNLLSLDDFD	1857
QY	1845	RVTRTEKIYDDHRKFTLRILYDQAGRPSSLWSPSSRLNGVNVYISPGGYIAGIQGIMSER	1904
Db	1858	RVTRTEKIYDDHRKFTLRILYDQAGRPSSLWSPSSRLNGVNVYISPGGYIAGIQGIMSER	1917
QY	1905	MEYDQAGRITSRIPADGKTWSYTYLEK---SMVLLHSQRQYIFEFDKNDRLSSVTMPN	1960
Db	1918	MEYDQAGRITSRIPADGKTWSYTYLEKAGVQSMVLLHSQRQYIFEFDKNDRLSSVTMPN	1977
QY	1961	VARQTLETIRSVGYRNIIYQPPPEGNASVIQDFTEDGHLLHTFYLGTGRRRVIYKYGKLSKL	2020
Db	1978	VARQTLETIRSVGYRNIIYQPPPEGNASVIQDFTEDGHLLHTFYLGTGRRRVIYKYGKLSKL	2037
QY	2021	AETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNAR	2080
Db	2038	AETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNAR	2097
QY	2081	FDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAAYMT	2140
Db	2098	FDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAAYMT	2157
QY	2141	HTKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNMGRVVVKELKVGYPYANTTRYSEYDADG	2200
Db	2158	HTKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNMGRVVVKELKVGYPYANTTRYSEYDADG	2217
QY	2201	QLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEDGF	2260
Db	2218	QLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEDGF	2277
QY	2261	LRQGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNP	2320
Db	2278	LRQGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNP	2337
QY	2321	TKVTHLYNHSSEITSLYYDLOQHLFAMELSSGDEFIACDNIGTPLAVFSGTGLMIKQI	2380

Db 2338 TKVTHLYNHSSEITSLYDYDQLQHLFAMELSSGDEFYIACDNIGITPLAVFSGTGLMIKQI 2397  
QY 2381 LYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHFWKHL 2440  
Db 2398 LYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHFWKHL 2457  
QY 2441 SSVNMPENLYMFKNPNPISNSQDIKCFMTDVNSWLLTFGQFQHNVPYKPKDMDAMERS 2500  
Db 2458 SSVNMPENLYMFKNPNPISNSQDIKCFMTDVNSWLLTFGQFQHNVPYKPKDMDAMERS 2517  
QY 2501 YELIHTQMKTOEWDNSK-----SILGVQCEVQKQKAFVTLERFDOLYGSTITSCQ 2552  
Db 2518 YELIHTQMKTOEWDNSKVIPAQCQSILGVQCEVQKQKAFVTLERFDOLYGSTITSCQ 2577  
QY 2553 APKTKKFASSGSVFGKGVKFKALDKGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTID 2612  
Db 2578 APKTKKFASSGSVFGKGVKFKALDKGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTID 2637  
QY 2613 GVDTHYFVKPGPSEGLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGA 2672  
Db 2638 GVDTHYFVKPGPSEGLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGA 2697  
QY 2673 LCLNTRYGTTLDEEKARVLELQRAVRQAWAREQQRLREGEGLRAWTEGEKQQVLSTG 2732  
Db 2698 LCLNTRYGTTLDEEKARVLELQRAVRQAWAREQQRLREGEGLRAWTEGEKQQVLSTG 2757  
QY 2733 RVQGYDGFVIVSQPELSDSANNIHFMRQSEMGR 2769  
Db 2758 RVQGYDGFVIVSQPELSDSANNIHFMRQSEMGR 2794

RESULT 7  
US-10-042-865-52  
; Sequence 52, Application US/10042865  
; Publication No. US20040029216A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Vernet, Corine A.M  
; APPLICANT: Taylor, Sarah  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Miller, Charles E  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Boldog, Ference L  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Rothenberg, Mark E  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Peyman, John  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Gunther, Erik  
; APPLICANT: Stone, David  
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of  
; TITLE OF INVENTION: Using the Same  
; FILE REFERENCE: 21402-537  
; CURRENT APPLICATION NUMBER: US/10/042,865  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: 60/260,417  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/260,831

; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: 60/272,338  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/274,876  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/284,704  
; PRIOR FILING DATE: 2001-04-18  
; NUMBER OF SEQ ID NOS: 264  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 52  
; LENGTH: 2715  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-042-865-52  
  
Query Match 68.7%; Score 10227; DB 12; Length 2715;  
Best Local Similarity 66.7%; Pred. No. 0;  
Matches 1864; Conservative 370; Mismatches 457; Indels 102; Gaps 19;  
  
QY 1 MDVKERKPYRSLTR-RRDAERRYTSSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57  
Db 1 MDVKERRPYCSLTKRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDYRLLYG 60  
  
QY 58 SRVKDIVPOEAEFCRTGANFTLRELGLEEVTTPHGTLYRTDIGLPQCYSNMGAGSDADM 117  
Db 61 NRVKDLVHREADEYTRQGNFTLRQLGVCESTRRGVAFCAEMGLPHRGYSISAGSDADT 120  
  
QY 118 EADTVLSPHPVRLWGRSTRSGRSSCLSSRANSNLITDTEHEN---TETDHPGGLQNH 174  
Db 121 ENEAVMSPEHAMRLWGRGVKSGRSSCLSSRANSALTITDTEHENRSDSESEPSNPGQP 180  
  
QY 175 RLRTPPPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQ--- 230  
Db 181 TLQPLPP--SHKQHPAQHH-PSITSLNRNSLTNRNQSPAP-----PAALPABLQTT 229  
  
QY 231 -EPAHAQENWLLNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHDGAYS DGHFLPKP 289  
Db 230 PESVQLQDSWVLGSNVPLESR-----HFLFKT 256  
  
QY 290 G-GTSPFLCTTSPGYPLTSTVYSPPPRPLPRSTFAFPANLKKPSKYCNWKAALSAIV 348  
Db 257 GTGTPLEFSTATPGYTMASGVSPPTRPLPRNTLSAFAFKKSSKYCSWRCALCAVG 316  
  
QY 349 ISATLVILLAYFVAMHLFGLNWHLOPMEGQMYE---ITEDTASSWPVPTDVSLYPSGGTG 405  
Db 317 VSVLLAILLSYFIAMHLFGLNWHLOQTENDTTFENGKNSDT-----VPTNTVSLPSG--- 368  
  
QY 406 LETPDRKKGKGTTEGKPSFFPEDSFIDSCEIDVGRRASQKIPPGTFWRISOVFIDHPVHLK 465  
Db 369 -----DNKLGGFTHENNTIDSGELDIGRRAIQEVPPGIFWRSOLFIDQPQLK 417  
  
QY 466 FNVSLGKAALVGIYGRKGLPPSHTQDFVELLDGRLLTQEARSLGTPRQSRGTVPSS 525  
Db 418 FNISLQKDALIGVYGRKGLPPSHTQYDFVELLDGSRLLIAREQNLVSEERAGRARVSL 477  
  
QY 526 HETGFIQYLDSGIWHLAFYNDGKSEWVSFLTIAESVDNCPNSCYNGDCTISGTHCFL 585  
Db 478 HEAGFIQYLDSGIWHLAFYNDGKNPEQVSFNTIVIESVVECPRNCHNGECVSGTCHCFF 537  
  
QY 586 GFLGPDGGRASCPVLCSGNGOYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTGT 645  
Db 538 GFLGPDGGRASCPVLCSGNGOYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTGT 597  
  
QY 646 CICNPGYKGESCEEVDMDPTCSGRGVCVRGECCHCFVWGWTNCTETPRATCLDQCSEHGT 705  
Db 598 CACNSGYKGENCEEADCLDPGCSNHGVCIHGECHCNPGWGSNCEILKTMCADQCSGHT 657  
  
QY 706 FLPTDTGLSCDPSWTGHDCSIETCAADCGGHGVCVGTGTCRCEDGMMGAACDQACHPRCA 765  
Db 658 YLQESGSGTCDPNWTGPDCSNEICSDVCGSHGVCMYGGSCRCRCEGTGPACNQRACHPRCA 717  
  
QY 766 EHGTCTRDGKCECSPGWNGEHCCTIAHYLDVRVWKEGCPGLCNGNRCRTLDLNGWHVCQLGW 825

Db 718 EHGTCCKGCECSQGMNGEHCTIAHYLDKIVKEGCPGLCNSNGRCLTDQNGWHVCVCPGW 777  
QY 826 RGAGCDTSMETACGSDKNDGDGLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVP 885  
Db 778 RGAGCDVAMETLCTDSKNEGDGLIDCMDPDCCLQSSCQNPYCRGLPDPQDIISQSLQT 837  
QY 886 VSQONLHSFYDRIKELVGRDSTHIIPGENPDCGHACVIRGOVMTSDGTPLVGVNISFVN 945  
Db 838 PSQQAASFYDRISFLIGSDSTHVLPGESPKNKSLASVIRGOVLTADGTPLIGVNVSLFH 897  
QY 946 NPLFGYTISRQDGSFVLVNGGSIILRFRERAPPITQEHTLWLWDRFPFVMTIIMRHEE 1005  
Db 898 YSEYGYTITRQDGMFDLVANGGASLTTLVERSPFLTQYHTVWIPWNVFYVMDTLVMKKEE 957  
QY 1006 NEIPSCDLNFPARNPVPVSPPLTSFASCAEKGPIVPEIQALQOEISISGCMRLSYLS 1065  
Db 958 NDIPSCDLGCFVRPSPPIIVSSPLSTFFRSPEDSPIIPETQVLHEETIPGTDLKLSYLS 1017  
QY 1066 SRTPGYKSVLRI SLTHPTTIPFNLKVKHLMVAVEGRLEFRKWFAPADLSYFIWDKTDVYN 1125  
Db 1018 SRAAGYKSVLKITMTQAVIPFNLKVKHLMVAVVGRFLQKWFAPASNLAYTFIWDKTDAYN 1077  
QY 1126 QKVFLGLSEAFVSQVGEYESCPDLILWEKRTVLQGEIDASKLGWSLDKHHALNIQSGI 1185  
Db 1078 QKVYGLSEAVVSQVGEYESCLDLTLWEKRTAVLQGYELDASNMGWTLDDKHVLDVQNGI 1137  
QY 1186 LHKNGENQFVSQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYV 1245  
Db 1138 LYKNGENQFISQPPFWSSIMGNRRRSISCPSCNGQADGNKLLAPVALACIGIDGSLYV 1197  
QY 1246 GDFNYIRRIFFPSGNVTNILELRNKDFRHSHPAKHYLATDPMGSAVFLSDSNSRRVPFKI 1305  
Db 1198 GDFNYVRRIFPSGNVTSVLELRNKDFRHSNPAHYLYLATDPTGDLYVSDTNTRRIRYP 1257  
QY 1306 KSTVVVKOLVKNSEWVAGTGDOCLPFDTRCGDGGKATEATLTNPRGITVDKFGLIYFVD 1365  
Db 1258 KSLTGAKDLTKNAEVVAGTGEQCLPFDEARCGDGGKAVEATLMSPKGMAIDKNGLIYFVD 1317  
QY 1366 GTMIRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVD 1425  
Db 1318 GTMIRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSLYVD 1377  
QY 1426 NNVLQISENHQVRIAVGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLVIAE 1485  
Db 1378 NNVLQITENPQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYSGVLVITE 1436  
QY 1486 TDEKKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDPFGDDGYAKDAKLNTSPSLAVC 1545  
Db 1437 TDEKKINRIRQVTTIDGEISLVAGIPSECDCKNDANCDYQSGDGYAKDAKLNAPSLLAAS 1496  
QY 1546 ADGELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYFDDTTGKHLTYQSLPTGD 1605  
Db 1497 PDGTLYIADLGNIRIRAVSKNKLNSMNFYEVASPTDQELYIFDINGTHQYTVSLVTGD 1556  
QY 1606 YLYNFTYTGDDITLITDNNGNMNVNRDSTGMPLWLVPDQVYVWVTMTGNSALKSVTT 1665  
Db 1557 YLYNFSYNDNDVTAVTDNNGNTLRIRDPNRMPVRVSPNQVWLITGTNGCLKSMTA 1616  
QY 1666 QGHELAMTYHNGSGLLATKSNENGTTFYEYDSFGRLTNTVFTPTGOVSSFRSDTSSVH 1725  
Db 1617 QGLELVLFTHGNSGLLATKSDETGWTTFDYDSEGRLTNTVFTPTGVVTLNHLGMDKAIT 1676  
QY 1726 VQVETSSK-DDVTITNLSASGAFYTLQDQVRNSYIIGADGSLRLLLLANGMEVALQTEP 1784  
Db 1677 VDISSSREEDVSITNSLSSIDSFTVMVQDLRNSYQIGYDGLRIFAYASGLDSHVQTEP 1736  
QY 1785 HLLAGTVNPTVGNKRNVTLPIDNGLNLVWRQKEQVTVFGRRLRVHNRNLLSLDFFD 1844  
Db 1737 HVLAGTANPTVAKRNMTLPGENGNLVWRFRKEQAQGVNFGKLRVNGRNLVSDVDFD 1796  
QY 1845 RVTRTEKIYDDHRKFTLRILYDQAGRPSPSLWSPSSRLANGVNVTVSPGGYIAGIQGIMSER 1904  
Db 1797 RTTKTEKIYDDHRKFLLRIAYDTSGHPTLWLPSSKLMAVNVTVSSGTQIASIQRTTSEK 1856

QY 1905 MEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFFDKNDRLSSVTMPNVARQ 1964  
Db 1857 VDYDSQGRIVSRVPADGKTWSYTYLEKSMVLLHLSQRQYIFEDVMDRLSAITMPSVARH 1916  
QY 1965 TLETIRSVGYRNIYQPPPEGNASVIQDFTEDGHLLHTFYLGTGRVIYKYGKLSKLAETL 2024  
Db 1917 TWQTIRSIGYRNINYPPEASNASIIITDNEEGLLLQTAFLGTSRRVLFKYRQTRLSEIL 1976  
QY 2025 YDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEEGMVARFDYN 2084  
Db 1977 YDSTRVSFTYDETAGVLTNVLQSDGFTCTIRYRQIGPLIDRQIFRFPSEDGMVARFDYS 2036  
QY 2085 YDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFKGFGVIYDINQIITTAVMTHTKH 2144  
Db 2037 YDNSFRVTSMQGVINETPLPIDLYQDFFDISGKVEQFKGFGVIYDINQIISTAVMTYTKH 2096  
QY 2145 FDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVKELKVGVPYANTTRYSYEYDADGLOQT 2204  
Db 2097 FDAHGRIKEIQYEIFRSLMYWITIYQDNMGRVTKREIKIGPFANTTKYAYEYDVGLOQT 2156  
QY 2205 VSINDKPLWRYSDNLGNLHLLSPGNSARLTPLRYDIRDRITELGDDVQYKMDDEGFLROR 2264  
Db 2157 VYLNEKIMWRYNDNLGNLHLLNPPSSARLTPLRYDLRDRITELGDDVQYRLDEDEGFLROR 2216  
QY 2265 GGDIFEYNSAGLLIKAYNRAGSWSVRYDGLGRRVSSKSSHHLQFFYADLTNPTKVT 2324  
Db 2217 GTEIFEYSSKGLLTRVYSGSGWTVIYRYDGLGRRVSSKSLGQHLQFFYADLTYPTRIT 2276  
QY 2325 HLYNHSSEITSLYYDLQHLFAMELSSGDEFYIACDNIGTFLAVFSGTGLMIKQILYTA 2384  
Db 2277 HVYNHSSEITSLYYDLQHLFAMELSSGDEFYIASDNTGTFLAVFSSNGLMLKQTOYTA 2336  
QY 2385 YGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMRDRDYDVLAGRWTSPDHELKHLSSNV 2444  
Db 2337 YGEIYFDSNVDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDIEIWKRI-GKDP 2395  
QY 2445 MPFNLYMFKNNPISNSQDIKCFMTDVSNSWLLTFFGQLHNVIPGYPKPDMDAMEPSYELI 2504  
Db 2396 APFNLYMFRNNNPASKIHVDKDYITDVSNSWLVTFGFHLHNAIPGFPVPKFDLTPEPSYELV 2455  
QY 2505 HTQMKTOEWDNSKSIILGVQCEVQKQLKAFVTLERFDQLYGSTITSCQQAQPKTKK----- 2558  
Db 2456 ---KSQWEDVPPIFGVQQQVQVQAKAFSLGKMAEV-----QVSRKAGAEQSW 2502  
QY 2559 --FASSGSVFGKGVKFKALDKGRVTTDIIISVANEDGRRVAAATLNHAHYLENLHFTIDGVD 2616  
Db 2503 LWFATVKSILIGKVMVLAVSQGRVQTNVLNIANEDCIKVAALNNAPYLENLHFTIEGKDT 2562  
QY 2617 HYFVKPGPSEGDAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDIOLQYCALCLN 2676  
Db 2563 HYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVVNGRTRRFADVEMQFGALAH 2622  
QY 2677 TRYGTTLDEEKARVLELQARQVQAWAREQORLREGEEGLRAWTEGEKQOVVLTGRVQG 2736  
Db 2623 VRYGMTLDEEKARILEQARQALARAWAREQORVRDGEAGARLWTEGEKQOVVLTGRVQG 2682  
QY 2737 YDGFVVISVEQYPELSDSANNHFMROSEMGR 2769  
Db 2683 YDGYVLSVEQYPELADSNANNIQLRQSEIGKR 2715

RESULT 8

US-10-029-020-51

; Sequence 51, Application US/10029020

; Publication No. US2004003971A1

; GENERAL INFORMATION:

; APPLICANT: Gangolli et al.

; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-225

; CURRENT APPLICATION NUMBER: US/10/029,020

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/256,704

; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: 60/311,590  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/257,314  
; PRIOR FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: 60/311,613  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/315,617  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/307,506  
; PRIOR FILING DATE: 2001-07-24  
; PRIOR APPLICATION NUMBER: 60/322,358  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/294,075  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: 60/288,153  
; PRIOR FILING DATE: 2001-05-02  
; NUMBER OF SEQ ID NOS: 190  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 51  
; LENGTH: 2715  
; TYPE: PR1  
; ORGANISM: Mus musculus  
; US-10-029-020-51

Query Match 68.7%; Score 10227; DB 12; Length 2715;  
Best Local Similarity 66.7%; Pred. NO. 0;  
Matches 1864; Conservative 370; Mismatches 457; Indels 102; Gaps 19;

QY	1	MDVKERKPYRSLTR-RRDAERRYTSSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG	57
Db	1	MDVKERRPYCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDYRLLYG	60
QY	58	SRVKDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADM	117
Db	61	NRVKDLVHREADEYTRQGNFTLRQLGVCESATRRGVAFCAEMGLPHRGYSISAGSDADT	120
QY	118	EADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHEN---TETDHPGGLQSHA	174
Db	121	ENEAVMSPEHAMRLWGRGVKSGRSSCLSSRSNSALTLDTEHENRSDSSEQPSNPGQP	180
QY	175	RLRTPPPPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDSLGSGEPPAGGAQ----	230
Db	181	TLQPLPP--SHKQHPAQHH-PSITSLNRNSLTNRNQSPAP-----PAALPAELQTT	229
QY	231	-EPAAQENWLLNSNIPLETRNLGKQPFGLGTLDQNLJEMDILGASRHDGAYSDDHFLFKP	289
Db	230	PESVQLQDSWVLSNVPLESR-----HFLFKT	256
QY	290	G-GTSPPLCTTSPGYPLTSSVYSPPPRPLPRSTFARPAFNLLKPKSKYCNWKCAALSAIV	348
Db	257	GTGTTPLFSTATPGYTMASGSVYSPPTPLPRNTLSRSFAFKFKSSKSKYCWRCALTCAVG	316
QY	349	ISATLVILLAYFVAMHLFGLNWHLQPMEGQMYE---ITEDTASSWPVPTDVSLYPSGGTG	405
Db	317	VSVLLAILLSYFIAMHLFGLNWHLQQTENDTFENGKVNSDT-----VPTNTVSLPSG---	368
QY	406	LETDPDRKGKGTTEGKPSFPFEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLK	465
Db	369	-----DNGKLGGFTHENNTIDSGELDIGRRAIQEVPPGIFWRSQVFIIDQPFQFLK	417
QY	466	FNVSIGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRSGRTVPPSS	525
Db	418	FNISLQKDALIGVYGRKGLPPSHTQYDFVELLDGSRLLIAREQNLNVESEAGRQARSVSL	477
QY	526	HETGFIQYLDSDGIWHLAFYNDGKESEVSWFLLTAIESVDNCPNSCYNGNDCISGTCCHFL	585
Db	478	HEAGFIQYLDSDGIWHLAFYNDGKNPEQVSFNTIVIESVVECPRNCHNGGECVSGTCHCFP	537
QY	586	GFLGPDCCRASCPLVCSGNGOYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTITGT	645
Db	538	GFLGPDCSRACPLVCSGNGQYSKGRCLCFSGWKGTEDCVPTTQCIDPQCGRGICIMGS	597

QY	646	CICNPGYKGESCEEVD CMDPTCSGRGVCVRGECHCFVGWGTNCETPRATCLDQC SGHT	705
Db	598	CACNSGYKGENCEEADCLDPCCSNHGVCIHGECHCNPGWGSNCEILKTMCADQCSGHGT	657
QY	706	FLPDTGLCSCDPSWTGHDCSIEICAADCGGHCVCVGGTCRCEDGWMGAACDQACHPRCA	765
Db	658	YLOESGSCCTDPNWTGPDSCNEICSVDCGSHGVCMGGSCREEGTGPACNQRACHPRCA	717
QY	766	BHGTCTRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCRTLNLNGHVCVQLGW	825
Db	718	BHGTCKDGKCECSQGWNGEHCTIAHYLDKI VKEGCPGLCNSNGRCRTLQNGHVCVQPGW	777
QY	826	RGAGCDTSMETACGDSKNDGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIQETQVP	885
Db	778	RGAGCDVAMETLCTDSKDNEDGLIDCMDPDCCLOQSSCONQPYCRGLPDPQDIISQSLQT	837
QY	886	VSQONLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFVN	945
Db	838	PSQAAKSFYDRISFLIGSDSTHVLPGESPENKSLASVIRGQVLTADGTPLIGVNVSFLLH	897
QY	946	NELFGYTI SRQDGSFDLVNTGGISIIILRFERAPFITQEH TLWLPWDRFFVME TIIMRHEE	1005
Db	898	YSEYGYTITRQDGMFDLVANGGASLT LVFERSPFLTQYHTVWIPWNVYVMDTLVMKKEE	957
QY	1006	NEIPSCDLSNFARPNPVPSPLTSFASSCAEKGPIVPEIQALQEEISISGCKMRLSYLS	1065
Db	958	NDIPSCDLSGFVRPSPIIVSSPLSTFFRSPSPEDSPIPETQVLHEETIPGTDLKL SYLS	1017
QY	1066	SRTPGYKSVLRISLTHPTIPFNLMKVHLMVAVEGRFRKWFAPFAAADLSYFFIWDKTDVYN	1125
Db	1018	SRAAGYKSVLKITMTQAVIPFNLMKVHLMVAVVGRLFKQKFFPASP NLAYTFIWDKTDAYN	1077
QY	1126	QKVFLSEAFVSUGYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKKHALNIQSGI	1185
Db	1078	QKVYGLSEAVSVUGYEYESCLDLTLWEKRTAVLQGYELDASNMGWTLDKHHVLDVQNGI	1137
QY	1186	LHKGNGENQFVSQOPPVIGSIMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYV	1245
Db	1138	LYKGNGENQFISQOPPVVSSIMGNGRRRSISCPSCNGQADGNKLLAPVALACGIDGSLYV	1197
QY	1246	GDFNYIRRIFFPSGNVTNILELRNKDFRHSHPAHKYVILATDPMMSGAVFLSDSNSRRVFKI	1305
Db	1198	GDFNYVRRIRFPNGVTSVLELRNKDFRHSNPAHRYVILATDPTGDLVYSDTNTFRIRYRP	1257
QY	1306	KSTVVVKDLVKNSEVVAGTGDQCLPPDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVD	1365
Db	1258	KSLTGAKDLTKNAEVVAGTGEQCLPFDEARCGDGGKAVEATL MSPKGMAIDKNGLIYFVD	1317
QY	1366	GTMIIRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLD	1425
Db	1318	GTMIIRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLD	1377
QY	1426	NNVVLQISENHQVRI VAGRPMHCQVPGIDHFLLSKVIAHATLESATALAVSHNGVLYIAE	1485
Db	1378	NNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYSGVLYITE	1436
QY	1486	TDEKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDCPSGDDGYAKADKLNTPSSLAVC	1545
Db	1437	TDEKINRIRQVTTDGEISLVAGIPSECDCKNDANCDCYQSGDGYAKADKLNAPSSLAAS	1496
QY	1546	ADGELYVADLGNIRIRFIRKNKPFPLNTQNMVELSSPIDQELIYFDTTGKHLYTQSLPTGD	1605
Db	1497	PDGTLTYIADLGNIRIRAVSKNKLNSMNFYEASPTDQELIYFIDINGTHQYTVSLVTGD	1556
QY	1606	YLYNFTYTGDDITLITDNNGNMNVNRRDSTGMPLWLVPDQVYVVTMTGNSALKSVTT	1665
Db	1557	YLYNFSYNDNDVTA VTDNSGNLIRIRDPNRMFVRVSPNQVIMLTIGTNGCLKSMTA	1616
QY	1666	QGHELAMWTHGNSGLLATKSNENGWTTTFEYEDSFGRLTNTVTFPTQVSSFSRSDTSSVH	1725
Db	1617	QGLELVLFTHGNSGLLATKSDETGTTFFDYDSEGRLTNTVTFPTGVVTLNHGMDKAIT	1676
QY	1726	VQVETSSK-DDVTITTNLSASGAFYTL LQDQVRNSYIIGADGSLRLLLLANGMEVALQTEP	1784

Db 1677 VDISSSREEDVSTISNLSSIDSFYTMVQDLRNSYQIGYDGLSIFAYASGLDSHYQTEP 1736  
QY 1785 HLLAGTVNPTVGRNVTLPLIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSLDFD 1844  
Db 1737 HVLAGTANPTVAKRNTLPGENGQNLVWRFRKEQAQGVNFGRKLVRNGRNLSSVDFD 1796  
QY 1845 RVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTVSPGGYIAGIQGIMSER 1904  
Db 1797 RTTKTEKIYDDHRKFLRLIAYDTSGHPTLWLPSSKLMVNVTVSYSTGTQIASIQRTTSEK 1856  
QY 1905 MEYDQAGRITSRIFADGKTWSYTYLEKSMVLLLSQRQYIFEDKNDRLSSVTMPNVARQ 1964  
Db 1857 VDYSQGRIVSRVPADGKTWSYTYLEKSMVLLLSQRQYIFEDMDWDRLSAITMPSVARH 1916  
QY 1965 TLETIRSVGYRNIYQPPPEGNASVIOQDFEDGHLHTFYLGTGRVVIYKYGKLSKLAETL 2024  
Db 1917 TMQTIRSIGYRNIYNPPESNASIITDYNEEGLLLQTAFLGTSRRVLFKYRRQTRLSEIL 1976  
QY 2025 YDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEEGMVNARFDYN 2084  
Db 1977 YDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLIDRQIFRSESDGMVNARFDYS 2036  
QY 2085 YDNSFRVTSQMAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAVMTHPKH 2144  
Db 2037 YDNSFRVTSQMGVINETPLPIDLYQDFDISGKVEQFGKFGVIYDINQIISTAVMTYTKH 2096  
QY 2145 FDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVVKELKVGYPANTTRYSYEYDADGQLQT 2204  
Db 2097 FDAHGRIKEIYEIFRSLMYWITIYQDNMGRVTKREIKIGPANTTKYAYEYDVGQLQT 2156  
QY 2205 VSINDKPLWRSYDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKXMDGDFLRQR 2264  
Db 2157 VYLNEKIMWRNYDLNGNLHLLNPSSSARLTPLRYDLDRITRLGDVQYRLDEDEGFLRQR 2216  
QY 2265 GGDIFEYNSAGLLIKAYNRAGSWSVRYRVDGLRRVSSKSSHHLQFFYADLTNPTKVT 2324  
Db 2217 GTEIFEYSSKGLLTRVYSKSGGWTVIYRVDGLRRVSSKTLGQHLQFFYADLTNPTRIT 2276  
QY 2325 HLYNHSSEITSLYYDLQGLHFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTA 2384  
Db 2277 HVYNHSSEITSLYYDLQGLHFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLKQTYTA 2336  
QY 2385 YGEIYMDTNPENFQIIGYHGGLYDPLTKLVHMGRDDYDLVLAGRWTSPDHXLWHLSSSNV 2444  
Db 2337 YGEIYFDSNVDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDIEIWKRI-GKDP 2395  
QY 2445 MPFNLYMFKNNPISNSQDIKCFMTDVNSWLLTTFGQLHNVIPGYPKPMDMAMEPSYELI 2504  
Db 2396 APFNLYMFRNNPASKIHDVKDYITDVNSWLVTFGFHLNNAIPGFPVPKFDLTPEPSYELV 2455  
QY 2505 HTQMKTOEWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSCQQAQTKK----- 2558  
Db 2456 ----KSQWEDVPPIFGVQQQVARQAKAFSLGKMAEV-----QVSRRKAGAEQSW 2502  
QY 2559 --FASGSGVFGKGVKFKALKDGRVTTDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDT 2616  
Db 2503 LWFAVTKSLICKGWLAVSQGRVQTNVLNANEDCIKVAAVLNNAFYLENLHFTIEGKDT 2562  
QY 2617 HYFVKPGSPSGDLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLN 2676  
Db 2563 HYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVNGRTRRFADVEMQFALALH 2622  
QY 2677 TRYGTTLDBEKARVLELQRAVRQAWAREQOORLREGEGELRAWTEGEXQQVLSTGRVQG 2736  
Db 2623 VRYGMTLDBEKARILEQARQALAFARAREQOORVRDGEGARLWTEGEXKQRLSAGKVQG 2682  
QY 2737 YDGFVVSVEQYPELSDSANNIHPMQSEMGR 2769  
Db 2683 YDGYVYLSVEQYPELADSDANNIQLRQSEIGKR 2715

US-10-038-854-38  
; Sequence 38, Application US/10038854  
; Publication No. US20040022781A1  
; GENERAL INFORMATION:  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Li, Li  
; APPLICANT: Wolenc, Adam R  
; APPLICANT: Vernet, Corine  
; APPLICANT: Eisen, Andrew J  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Gorman, Linda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gusev, Vladimир Y  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Guo, Xiaojia S  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Erik  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Millet, Isabelle  
; APPLICANT: MacDougall, John R  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-230  
; CURRENT APPLICATION NUMBER: US/10/038,854  
; PRIOR FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: 60/258,928  
; PRIOR FILING DATE: 2000-12-29  
; PRIOR APPLICATION NUMBER: 60/259,415  
; PRIOR FILING DATE: 2001-01-02  
; PRIOR APPLICATION NUMBER: 60/259,785  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 60/269,814  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/279,832  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,833  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/279,863  
; PRIOR FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/283,889  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 60/284,447  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/286,683  
; PRIOR FILING DATE: 2001-04-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 411  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 38  
; LENGTH: 2721  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-038-854-38

Query Match 68.7%; Score 10221; DB 16; Length 2721;  
Best Local Similarity 67.0%; Pred. No. 0;  
Matches 1873; Conservative 362; Mismatches 461; Indels 98; Gaps 20;

QY 1 MDVKRRPKYRSLTR-RRDAERRYSSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57  
Db 1 MDVKERRPYCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDSSRLYG 60  
QY 58 SRVKDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADM 117

Db 61 NRKDLVHREADEFTROQNFTRLQGVCEPATRRGLAFCAEMGLPHRGYSISAGSDADT 120  
QY 118 EADTVLSPHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEHN--TETDHPGGLQWHA 174  
Db 121 ENEAVMSPHAMRLWGRGVKSGRSSCLSSRSNSALTTLTDTHEHNKSDSENEQPASNQOS 180  
QY 175 RLRTPPPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDSLSGEPPAGGAQ---- 230  
Db 181 TLQPLPP--SHQHSQHH-PSITSLNPNLSLTNRNQSPAP-----PAALPAELQTT 229  
QY 231 -EPAHAQENWLLNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHDGAYSDBGHFLFKP 289  
Db 230 PESVOLQDSWVLGSNVPLESR-----HFLFKT 256  
QY 290 G-GTSPFLCTTSPGYPLTSSTVSYSPPPRPLPRSTFARPAFNLKKPSKYCNWKCAALSIV 348  
Db 257 GTGTTPLFSTATPGYTMASGSVYSPPTRLPRNTLSRSAFKFKSKSKYCSWKCTALCAVG 316  
QY 349 ISATLVILLAYFVAMHLFGLNWHLOPMEGOMYE---ITEDTASSWPVPTDVSLYPSSGGTG 405  
Db 317 VSVLLAILLSYFIAMHLFGLNWLOQTENDTFENGKVNSDT-----MPINTVSLPSG--- 368  
QY 406 LETPDRKGKGTTEGKPSSEFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHPVHLK 465  
Db 369 -----DNGKLGFTQENNTIDSGELDIGRRAIQEIIPPGIFWRSQLFIDQPQFLK 417  
QY 466 FNVSLGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTOEARSLEGTFPRQSRGTVPSS 525  
Db 418 FNISLQKDALIGVYGRKGLPPSHTQYDFVELLDGSRLLIAREQSRLLETERAGRQARSVSL 477  
QY 526 HETGFIQYLDSGIWHLAFYNDGKESVVSFLTTAIESVDNCPNSCYGNGBDCISGTCHCFL 585  
Db 478 HEAGFIQYLDSGIWHLAFYNDGKNAEQVSFNTIVIESVVECPRNCHNGECVSGTCHCFP 537  
QY 586 GFLGPDCCGRASCPVLCNGQYMKGRCLCHSGWKAECDVPTNQCIDVACSNHGTCITGT 645  
Db 538 GFLGPDCSRACPVLCNGQYKGRCLCFSGWKTECDVPTTQCIDPQCGRGICIMGS 597  
QY 646 CIGNPGYKGESCEEVDMDPTCSGRGVCVRGECHECFVGMGGTNCETPRATCLDQCSEGHGT 705  
Db 598 CACSSGYKGESCEEADCDPCCSNHGVCIHGECHCSPGWSNCEILKTMCPDQCSEGHGT 657  
QY 706 FLPDTGLCSDPSWTGHDCSIIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCA 765  
Db 658 YLQESGSCTCDPNWTGPDSCNEICSVDCSGSHGVCWGGTCRCBEGWTGPACNQACHPRCA 717  
QY 766 EHGTCTRDGKCECSPGWNGEHCTIAHYLDRVV-----KEGCPGLCNGNGRCTLDLNGWHC 819  
Db 718 EHGTCKDGKCECSQGWNGEHCTIAHYLDKIVKDIXGKYGECPCPLCNSNGRCTLDQNGGHC 777  
QY 820 VCQLGWRGAGCDTSMETACGDSKDNDDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDII 879  
Db 778 VCQPGWRGAGCDVAMETLCTDSKDNEDGLIDCMDPDCCLOQSSCQNPYCRGLPDPQDII 837  
QY 880 QETQVPVSQQNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGV 939  
Db 838 SOSLQSPSQAAKSFYDRISFLIGSDSTHVIIPGESPFNKSLASVIRGQVLTADGTPLIGV 897  
QY 940 NISFVNPLFGYTISRQDGSFDLVNTNGGISIIILRFERAPFITQHTLWLPWDRFFVMETI 999  
Db 898 NVSFFHYPEYGYTITRQDGMFDLVANGGASLTLVFERSPFLTQYHTVWIPMNVFYVMDTL 957  
QY 1000 IMRHEENIPSCDLSNFAARNPVVSPSLTSFASSCAEKGPVPEIQALQEEISISGCKM 1059  
Db 958 VMEKEENDIPSCDLSGFVRPNPIVSSPLSTFFRSSPEDSPIIPETQVLHEETIIPGTDL 1017  
QY 1060 RLSYLSRTPGYKSVLRISLTHPTIPFNLMKVHLMVAVEGSLFRKWFAAAPDLSYFIWD 1119  
Db 1018 KLSYLSRAAGYKSVLKITWTQSIIPFNLMKVHLMVAVVGRILFQKWFASPANLAYFIWD 1077  
QY 1120 KTDVYNQKVGFLSEAFVSVGYEYESCPDILILWEKRTTVLQGYEIDASKLGWSLDKXHAL 1179

Db 1078 KTDAYNQKVYGLSEAVSVGVYEYESCLDLTLWEKRTAILQGYELDASNMGWTLDKXHLV 1137  
QY 1180 NIOSGILHKNGENQFVSQQPPVIGSIMGNGRRRSISCPSCNGLADGNKLLAPVALTCGS 1239  
Db 1138 DVQNGILYKNGENQFISQQPPVWSSIMGNGRRRSISCPSCNGQADGNKLLAPVALACGI 1197  
QY 1240 DGSLYVGDFNYIRIFPSPGNVTNILELRNKDPRHSHSPAHHKYLLATDPMGAVFLSDSNS 1299  
Db 1198 DGSLYVGDFNYIRIFPSPGNVTSVLELRNKDPRHSSNPAHRYLLATDPVTGDLVSDTNT 1257  
QY 1300 RRVFYKIKSTVVVKDLVKNSEWAGTGDOCLPDDTRCGDGGKATEATLTPRGITVDDKFG 1359  
Db 1258 RRIYRPKSLTGAKDLTKNAEVVAGTGEQCLPDEARCGDGGKAVEATLMSPKGMAVDKNG 1317  
QY 1360 LIYFVDTGTMIRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDN 1419  
Db 1318 LIYFVDTGTMIRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDN 1377  
QY 1420 SLYVLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNG 1479  
Db 1378 SIYVLDNNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYSG 1436  
QY 1480 VLYIAETDEKKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNT 1539  
Db 1437 VLYITETDEKKINRIRQVTTDGEISLVAGIPSECDCKNDANCDGSGDDGYAKDAKL 1496  
QY 1540 SSIYVLTGADGELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFTTGKHL 1599  
Db 1497 SSIYVLTGADGELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFTTGKHL 1556  
QY 1600 SLPTGDLYNFTYTGDDITLITDNNGNMNVNRDSTGMPLWLVPDQGVYVWMTGNTSA 1659  
Db 1557 SLVTDGDLYNFYSNDNDITAVTDSNGNTLIRRRDNRMPVRVWSPDNQVIMLTIGTNGC 1616  
QY 1660 LKSVTTQGHHELAMTYHNGSGLLATKSNENGTTFYVYDSFGRLLTNVTFPTQGVSSFRSD 1719  
Db 1617 LKGMTAQGLELVLFYTHGNSGLLATKSDSETGTTTFDYDSEGRLLTNVTFPTQGVV 1676  
QY 1720 TDSSVHVQVETSSK-DDVTITNLSASGAFYTLTQDVNRNSYVIGADGSLRLLLANGMEV 1778  
Db 1677 MDKAITVDIESSSREEDVSTSNLSSIDSFYTMVQDLRNSYQIGYDGLRIIVASGLDS 1736  
QY 1779 ALQTEPHLLAGTVNPTVGKRVNTPIDNGLNLVWRQKQEQARGQVTVFGRRLRVHNRNL 1838  
Db 1737 HYQTEPHVLAGTANPTVAKRNMTLPGENGNLVWRFRKQEQAGKVNFGKRLVRN 1796  
QY 1839 LSLDFRVRTRTEKIYDDHRKFTLIRILYDQAGRPSLWSPSSRLNGVNVYSPGGYIAGIOR 1898  
Db 1797 LSVDFRRTTKTEKIYDDHRKFTLIRIAYDTSGHPTLWLPSSKLMVAVNVYSSGTQIASIOR 1856  
QY 1899 GIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLLSSVTM 1958  
Db 1857 GTTSEKVDYDGGQGRIVSRVFPADGKTWSYTYLEKSMVLLHLSQRQYIFEYDMWDRLSA 1916  
QY 1959 PNVARQTLTETIRSVGYRNIYQPPENASVIODFETEDGHLHFTFYLGTRRRVIYKYGKLS 2018  
Db 1917 PSVARHTMQTIRSIGYRNIYNPPESNASIITDYNEEGLLQTAFLGTSRRVLFKYRROT 1976  
QY 2019 KLAETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQICPLIDRQIFRFTTEEGMVN 2078  
Db 1977 RLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQICPLIDRQIFRSEEDGMVN 2036  
QY 2079 ARFDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQKFGVYIYDINQIITAV 2138  
Db 2037 ARFDYSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGKFGVYIYDINQIISTAV 2096  
QY 2139 MTHTKHFDAYGRMKEVQYEIFRSLMYMVTQYDNMGRVVVKELKVGYPYANTTRYSEYDA 2198  
Db 2097 MTYTKHFDAGHRIKEIQYEIFRSLMYWITIYQDNMGRVTKREIKIGPFANTTKYAYEYDV 2156  
QY 2199 DGOLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYXMD 2258  
Db 2157 DGOLQTVYLNEKIMWRYNYDLNGLNHLNPSNSARLTPLRYDLRDRITRLGDVQYRLDED 2216

QY	2259	GFLRQGGDIF	EYNSAGLLIKAYN	RAGSWSVR	YRYDGLGRR	VSSKSSHHLQFF	YADLT	2311					
Db	2217	GFLRQGT	EIFEYSSKGLL	TRVYSKGS	GWTVIYRYD	GLGRRVSSK	TSLGQHLQFF	YADLT	2276				
QY	2319	NPTKVTHLYN	HSSEITSLY	DLQGHLFAME	LSGDEFIAC	DNIGTPLAV	FSGGLMIK	2378					
Db	2277	YPTRI	THVYNHSSE	ITSLYDLQGH	LFAEISSG	DEFIASN	TGTPLAVF	SNGMLK	2336				
QY	2379	QILYTAYGE	IYMDTNPFO	IIGYHGGLY	DPLTKLVHM	GRDYDVL	AGRWTSP	DPHELWKH	2438				
Db	2337	QIQYTAYGE	IYFDSNID	QLVIGFHGGLY	DPLTKLIHF	GERDYD	ILAGRWT	TPDIEIWK	2396				
QY	2439	LSSNVMPFN	LYMFKNNP	ISNSODIKCF	MTDVNSWLL	TGFGQLHN	VI	PGYKPDMD	AME	2498			
Db	2397	I-GKDPAP	FNLYMFRN	NNPASKIH	DVKYITDV	NSWLVTF	GFLHNAI	PGFPVPK	FDLTE	2455			
QY	2499	PSYELIHTQ	MKTQEW	NSKSI	LGVCQVQ	KLKAFTL	ERFDQ	LYGSTITS	CQQA	PKTKK	2558		
Db	2456	PSYELV---	KSQWDD	IPPIFGV	QQQVARQ	AKAFSLG	KMAEVQ---	VSRRRAG	GAQS	2507			
QY	2559	---	FASSGSV	FGKVKF	ALKDGRV	TTDII	SVANED	GRRVAA	ILNHA	YLENLH	FTIDGVD	2615	
Db	2508	WLWFATV	KS	LIGKGV	MLAVSQ	RGVQTN	VLNIANED	CIKVA	AVLNNA	FYLENLH	FTIEGKD	2567	
QY	2616	THYFVKP	GPSEGL	LA	ILGLSG	RRRTLE	NGVNVTV	SQINTV	LN	GRTRRY	TDIQLQY	GALCL	2675
Db	2568	THYFIK	TTTTPE	SLDGL	RLTSGR	KALENG	INVTVS	QSTTVV	NGRTRR	RFADVEM	QFGALAL	2627	
QY	2676	NTRYG	TTLDEE	KARV	ELARQ	RAVRQ	AWAREQ	ORLREGE	EGLRAW	TEGEKQ	QVLS	TGRVQ	2735
Db	2628	HVRYG	MTLDEE	KARILE	QARQ	ALARAR	AWAREQ	QRVRD	GEGARL	WTEGEK	RLLSAG	KVQ	2687
QY	2736	GYDGF	FVIS	VEQY	PELSD	SANNI	HFM	RQSE	MGR	2769			
Db	2688	GYDGY	VLSVEQ	YELAD	SANNI	QFLR	OSIG	R	2721				

```

; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 2725
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-038-854-36

Query Match      68.5%; Score 10201; DB 16; Length 2725;
Best local Similarity 66.9%; Pred. No. 0;
Matches 1872; Conservative 362; Mismatches 462; Indels 102; Gaps 21;

QY      1 MDVKERKPYRSLTP-RRDAERRYTTSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      1 MDVKERRYPCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDSSRLLYG 60

```

REFERENCE: 2104 250  
CURRENT APPLICATION NUMBER: US/10/038,854





Db 928 YLSRAAGYKSVLKITMTQSIIPFNLKMHLMVAVVGRLEFQKWFPASPNLAYTFIWDKTD 987  
Qy 1123 VYNQVFGLSBAFVSUGYEBESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHHALNIQ 1182  
Db 988 AYNQVYGLSEAVVSUGYEBESCLDLTLWEKRTAILQGYELDASNMGGWTLDKHHVLDVQ 1047  
Qy 1183 SGILHKGNGENQFVSQQPPVICSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGS 1242  
Db 1048 NGILYKNGENQFISQQPPVSSIMGNRRRSISCPSCNQADGNKLLAPVALACGIDGS 1107  
Qy 1243 LYVGDFNYIRRIFFPSGNVTNILELRNKDFRSHSHPAHKYIYLATDPMGAVFLSDSNSRRV 1302  
Db 1108 LYVGDFNYVRRIFPSGNVTSVLELRNKDFRSHSNPAHYIYLATDPVTGDLVYVSDTNTRR 1167  
Qy 1303 FKJFKSTVVVKOLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIY 1362  
Db 1168 YRPKSLTGAKDLTKNAEUVVAGTGEQCLPFDEARCGDGGKAVEATLMSPKGMAVDKNGLIY 1227  
Qy 1363 FVDFMIRRIDQNGIISTLLGNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLY 1422  
Db 1228 FVDGFMIRKVDQNGIISTLLGNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIY 1287  
Qy 1423 VLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLY 1482  
Db 1288 VLDNNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYGVLY 1346  
Qy 1483 IAETDEKINRIROVTTISGEISILVAGAPSGCDCKNANDCFSGDDGYAKDAKLNTPSSL 1542  
Db 1347 ITETDEKINRIROVTTDGEISILVAGIPSECDCKNANDCYQSGDGYAKDAKLAPSSL 1406  
Qy 1543 AVCADGELYVADLGNIRIRFIRKNKPFPLNTQNMVELSSPIDQELYLFDTTGKHLYTQSLP 1602  
Db 1407 AASPDGTLYIADLGNIRIRAVSKNKPLLNSMNFYEVASTDQELYIFDINGTHQYTVSLV 1466  
Qy 1603 TGDYLYNFYTYTGDGDTILITDNNGMVNVRRDSTGMPMLVVPDGOVYVWVTMGNTSALKS 1662  
Db 1467 TGDYLYNFSYSNDNDITAVTDSNGNTLRIRRDPNRMFVRVSPDNQVILWTIGTNGCLKG 1526  
Qy 1663 VTTQGHELAMNTYHGNSGLLATKSNENGWTTTFEYDYSFGRLTNTVTFPTGOVSSFRSDTDS 1722  
Db 1527 MTAQGLELVLFYTHGNSGLLATKSDTGTWTTFFDYDSEGRLTNTVTFPTGVVTNLHGDMDK 1586  
Qy 1723 SVHVQVETSSK-DDVTITTNLSASGAFYTLLODQVRNSYIIGADGSLRLLLLANGMEVALQ 1781  
Db 1587 AITVDIESSSREEDVSITSNLSSIDSFYTMVQDQLRNSYQIGYDGLSLRIIYASGLDSHYQ 1646  
Qy 1782 TEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEWRQRKEAQGVTVFGRRLRVHNRNLLSL 1841  
Db 1647 TEPHVLAGTANPTVAKRNMTLPGENGQNLVEWRFRKEAQGVKNVFGKLRVNGRNLLSV 1706  
Qy 1842 DFDRTVTEKIYDDHRKFTLRILYDQAGRPSSLWSPSSRLNGVNVITYSPGGYIAGIQRGIM 1901  
Db 1707 DFDRTTTEKIYDDHRKFLRLRIAYDTSGHPTLWLPSSKLMAVNVITYSSTGQIASIQRGTT 1766  
Qy 1902 SERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLLHSQROYIFEFDKNDRLSSVTMPNV 1961  
Db 1767 SEKVDYDQGRIVSRVFADGKTWSYTYLEKSMVLLLHSQROYIFEYDMWMDRLSAITMPSV 1826  
Qy 1962 ARQTTLETIRSVGYRNIYQOPPEGNASVIOQDFTEDGHLHHTFVLGTGRRVIYKYGKLSKLA 2021  
Db 1827 ARHTMQTIRSIGYRNIYNPPESNASIITDYNEEGLLLOTAFLGTSRRVLFKYRRQTRLS 1886  
Qy 2022 ETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEEGMVNARF 2081  
Db 1887 EILYDSIRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLIDRQIFRFSGEDGMVNARF 1946  
Qy 2082 DYNYNDSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFGFVYIYDINQIITTAVMTH 2141  
Db 1947 DYSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGFVYIYDINQIISTAVMTY 2006  
Qy 2142 TKHFDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVKKELKVGPPYANTTRYSYEYDADGQ 2201

Db 2007 TKHFDAGRIKEIQYEIFRSLMYWITIQYDNMGRVTXREIKIGPPANTTKYAYEYVDVGQ 2066  
Qy 2202 LQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYXMDDEGFL 2261  
Db 2067 LQTVYLNKIMWRYNYDLNGNLHLLNPSNSARLTPLRYDLDRITRLGDVQYRLDEGFL 2126  
Qy 2262 RQRGGDIFEYNSAGLLIKAYNEAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPT 2321  
Db 2127 RQRGTEIFEYSSKGLLTRVYSKSGWTVIYRYDGLGRRVSSKTSLGQHLQFFYADLTPT 2186  
Qy 2322 KVTHLYNHSSEITSLYYDLQGLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQIL 2381  
Db 2187 RITHVYNHSSEITSLYYDLQGLFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLKQIQ 2246  
Qy 2382 YTAYGEIYMDTNPNFQIIIGVHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKHLSS 2441  
Db 2247 YTAYGEIYFDSNIDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDIEIWKRI-G 2305  
Qy 2442 SNVMPFNLYMFKNNNPISNSQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPMDAMERSY 2501  
Db 2306 KDPAPFNLYMERNNPASKIHDVKDYITDVNSWLVTTFGFHLHNAIPGFVPKFDLTERSY 2365  
Qy 2502 ELIHTQMKTOEWDNSKSILGVQCEVQKQLKAFVTILERFDQLYGSTITSCQQAPKTK--- 2558  
Db 2366 ELV---XSQQWDDIPPIFGVQQQVARQAKAFLSLGKMAEVQ---VSRRRAGGAQSMLW 2417  
Qy 2559 FASSGSVFGKGVKFPALKDGRVTTDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDTY 2618  
Db 2418 FATVKSILICKGVMLAVSQGRVQTVNLNIANEDCIKVAAVLNNAPYLENLHFTIEGKDTY 2477  
Qy 2619 FVKPGPSEGLAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRYTDIQLQYGALCLNTR 2678  
Db 2478 FIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVVNGRTRRFADVEMQFGALALHVR 2537  
Qy 2679 YGTTLDEEKARVLELARQRAVRQAWAREQORLREGEEGLRAWTEGEKQQVVLSTGRVQGYD 2738  
Db 2538 YGMTLDEEKARILEQARQARALARAWAREQORVRDGEGARLWTEGEKROLLSAGKVQGYD 2597  
Qy 2739 GFFVISVEQYPELSDSANNIHFMRQSEMGR 2769  
Db 2598 GYYVLSVEQYPELADSANNIQFLRQSEIGRR 2628

RESULT 12  
US-10-038-854-42  
; Sequence 42, Application US/10038854  
; Publication No. US20040022781A1  
; GENERAL INFORMATION:  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Li, Li  
; APPLICANT: Wolenc, Adam R  
; APPLICANT: Vernet, Corine  
; APPLICANT: Eisen, Andrew J  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Gorman, Linda  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Guo, Xiaojia S  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Erik  
; APPLICANT: Smithson, Glennnda

APPLICANT: Millet, Isabelle  
APPLICANT: MacDougall, John R  
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-230  
CURRENT APPLICATION NUMBER: US/10/038,854  
CURRENT FILING DATE: 2003-01-22

PRIOR APPLICATION NUMBER: 60/258,928  
PRIOR FILING DATE: 2000-12-29

PRIOR APPLICATION NUMBER: 60/259,415  
PRIOR FILING DATE: 2001-01-02

PRIOR APPLICATION NUMBER: 60/259,785  
PRIOR FILING DATE: 2001-01-04

PRIOR APPLICATION NUMBER: 60/269,814  
PRIOR FILING DATE: 2001-02-20

PRIOR APPLICATION NUMBER: 60/279,832  
PRIOR FILING DATE: 2001-03-29

PRIOR APPLICATION NUMBER: 60/279,833  
PRIOR FILING DATE: 2001-03-29

PRIOR APPLICATION NUMBER: 60/279,863  
PRIOR FILING DATE: 2001-03-29

PRIOR APPLICATION NUMBER: 60/283,889  
PRIOR FILING DATE: 2001-04-13

PRIOR APPLICATION NUMBER: 60/284,447  
PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: 60/286,683  
PRIOR FILING DATE: 2001-04-25

Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 411

SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 42

LENGTH: 2613  
TYPE: PRT

ORGANISM: Homo sapiens  
US-10-038-854-42

Query Match 66.2%; Score 9856; DB 16; Length 2613;  
Best Local Similarity 65.2%; Pred. No. 0;  
Matches 1815; Conservative 345; Mismatches 437; Indels 188; Gaps 21;

1 MDVKERKPYRSLTR-RRDAERRYTSSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57  
1 MDVKERRYPSLTCSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDSSRLLYG 60  
58 SRVKDIVPQEAEEFCRTGANFTLRELGLBEVTPPHGTLRYTDIGLPQCGYSMGAGSDADM 117  
61 NRVKDLVHREADEPTR-----LQ----- 76  
118 EADTVLSPEHPVRLWGRSTRSGRSSCLSRANSNLTLTDTEHENTETDHPGGLQNHARLR 177  
77 -----QECPASNQCGST-----LQ----- 90  
178 TPPPPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDLHSLSGEPPAGGAQ-----EP 232  
91 --PLPPSHKQHSQAHH-PSITSLNRNSLTNRNQSPAP-----PAALPAELQITPES 139  
233 AHAQENWLLNSNIPLETNRNLGKQPFGLTQDNLIEMDILGASRHDGAYSDDGHFLFKPG-G 291  
140 VQLQDSWVLGSNVPLESR-----HFLFKTGTG 166  
292 TSPLEFCTTSPGVPPLTSSTVSPPPRPLPRSTFARPAFNLLKPKSKYCNWKCAALSAIVISA 351  
167 TPLFSTATPGYTMASGSVSPPTPLPRNTLSRSFAKFKKSKYCSWKCTALCAVGVSU 226  
352 TLVILLAYFVAMHLFGLNWHLOFMEQOMYE---ITEDTASSWPVPTDVSLYPSGGTGLET 408  
227 LLAILLSYFIAMHLFGLNWLQQTENDTFENGKVNSDT-----MPTNTVSLPSG----- 275  
409 PDRKKGTTTEKPSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSSQVFIHPVHLKFN 468  
276 -----DNKGLGGFTQENNTIDSGLDGRRAIQEIIPPGIFWRSSQLFIDQPQFLKFN 327  
469 SLGKAALVGIYGRKGLPPSHQTQDFVELLDGRRLLTQEARSLLEGTPRQSRGTVPSPSHET 528

328 SLQKDALIGVYGRKGLPPSHTQYDFVELLDGSRLLIAREQSRLLLETERAGRQARSVSLHEA 387  
529 GFIOYLDSGIWHLAFYNDGKESEVSVFLTTTALIESVDNCPNSCYNGNDCISGTCHCFLGFL 588  
388 GFIOYLDSGIWHLAFYNDGKNAEQVSNFTIVIESVVECPRNCHNGNCEVSGTCHCFPGFL 447  
589 GPDGGRASCPVLCSGNGQYMKGRCLCHSGWKAECVPTNQCIDVACSNHGTCTITGTIC 648  
448 GPDCSRACPVLCSGNGQYSGRCLCFSGWKGTECDVPTTQCIDPQCGRGICIMGSCAC 507  
649 NPGYKGESCEEVDMDPTCSGRGVVGRGCHGVGWTNCEPRATCQCSGHGTFPL 708  
508 NSGYKGESCEEADCDIPGCSNHGVCIHGECHCSPGWGSNCEILKTMCPDQCSGHGTYLQ 567  
709 DTGLCSDPSWTGDCSIEICAADCGHGVVCGGTCTCEDGWMGAACDQACHPRCAEHG 768  
568 ESGSCTDPNWTGDCSNEICSVDCSGHGVCGGTCTCEEGTGTCTNQRACHPRCAEHG 627  
769 TCRDGKCECSPGWNHGHCTIAHYLDRVVKEGCPGLCNGNCRCTLDLNGWHVCVQLGWRGA 828  
628 TCKDGKCECSHGWNHGHCTI-----EGCPGLCNSNGRCTLQNGWHVCVQPGWRGA 678  
829 GCDTSMETACGSDKNDGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQ 888  
679 GCDVAMETLCTDSKDNEDGLDCMDPDCCLOSSCONQPYCRGLPPQDIISQSLQSPSQ 738  
889 QNLHSFYDRIKFLVGRDSTHIIIPGENPFDGHCACVIRGOVMTSDGTPLVGVNIFVNNPL 948  
739 QAAKSFYDRISFLIGSDSTHVIIPGESPFNKSLASVIRGOVLTADGTPLIGVNVFFHYPE 798  
949 FGYTISRQDGSFDLVTNGGSIILRFRERAPFITQEHTLWLPWDRFVWETIIMRHEENEI 1008  
799 YGYTITRQDGMFDLVANGGASLTLPERSPFLTYHTVWIPWNVFVMDTLVMEKEENDI 858  
1009 PSCDLSNFARPNPVVSPSPLTSPASSCAEKGPVIVPEIQALQEEIISGCKMRLSYLSRT 1068  
859 PSCDLSGFVRPNPIIVSSPLSTFFRSPSPEDSPIIPETQVLHEETIPGTDLKLSSRA 918  
1069 PGYKSVLRISLTHPTIPFNLKMHLMVAVEGRFLFRKFAAAPDLISYFIWKTQVYNQKV 1128  
919 AGYKSVLKITMTQSIIPFNLKMHLMVAVVGRLFKQWFPASPFLAYTFIWKTDAYNQKV 978  
1129 FGLSEAFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKLGGWSDKHALNIQSGLHK 1188  
979 YGLSEAVSVGYEYESCLDLTLWEKRTAILQGYELDASNMGWGTLDKHVLDVQNGILYK 1038  
1189 NGNGNQFVSQPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDF 1248  
1039 NGNGNQFISQPPVIVSSIMNGRRRSISCPSCNGQADGNKLLAPVALACGIDGSLYVGDF 1098  
1249 NYIRRIFPNGVNTNILELRNKDFRHSHPAHKYIYATDPMGSAVFLSDSNRRVFKIKST 1308  
1099 NYVRRIFPNSGNTSVLELRNKDFRHSNPAHYIYATDPTGDIYVSDTNRRIYRPKSL 1158  
1309 VVVKDLVKNSEVAGTGDQCLPDDTRCGDGGKATEATLTNPRGITVDKFLIYFVDGTM 1368  
1159 TGAKDLTKNAEVVAGTGEQCLPFDEARCGDGGKAVEATLMSPKGMAVDKNGLIYFVDGTM 1218  
1369 IRRIDQNGIITLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNV 1428  
1219 IRKVDQNGIITLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDNNV 1278  
1429 VLOISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDE 1488  
1279 VLOITENRQVRIAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYSGVLYITETDE 1337  
1489 KKNIRIRQVTSGEISLVAGAPSGCDCKNDANDCDFSGDDGYAKDAKLTNPSSLAVCADG 1548  
1338 KKNIRIRQVTTDGEISLVAGIPSECDCKNDANDCDYQSGDGYAKDAKLSAPSSLAASPDG 1397  
1549 ELYVADLGNIRIRFIRKPKPFLNTQNMVELSSPIDQELYFDTTGKHLTYQSLPTGDYLY 1608  
1398 TLYIADLGNIRIRAVSKNPKPLNSMNFYEVASPTDQELYIFDINGTHQYTVSLVTGDYLY 1457





Db 2555 QOTERHNOAFMALE-----GQVISKRLHASIREKAGHWFATSTPIICKGIMFAVKGR 2608

Qy 2579 VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGD LAILGLSGGR 2638

Db 2609 VTTGISSATDDSRKIASVLSNAHYLEKMHYSIEGKDTHYFVKIGSADSDLVTLAMTSGR 2668

Qy 2639 RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYG--TTLDDEKARVLELAR 2695

Db 2669 KVLDSGVNVTVSQPTLLINGRTRRFTNIEFQYSTLLINIRYGLTADTLDEEKARVLDQAR 2728

Qy 2696 QRAVRQAWAREQQRLREGEGELRAWTEGEKQQVLSTGRVQGYDGFVFISVEQYPELSDSA 2755

Db 2729 QRALGSAWAKEQQKARDGREGSRVWTDGEKQQLLNTGRVQGYEGYVYLPVEQYPELADSS 2788

Qy 2756 NNHFMRQSEMGR 2769

Db 2789 SNIQFLRQNMGR 2802

RESULT 14

US-09-800-198-69

; Sequence 69, Application US/09800198

; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 69

; LENGTH: 2802

; TYPE: PRT

; ORGANISM: Gallus gallus

US-09-800-198-69

Query Match 64.6%; Score 9616.5; DB 10; Length 2802;

Best Local Similarity 61.3%; Pred. No. 0;

Matches 1775; Conservative 411; Mismatches 491; Indels 217; Gaps 27;

Qy 1 MDVKERKPYRSLTRRR-DAERRYTSSSADSEEGKAP-QKSYSSSETLKAYDQDARLAYGS 58

Db 1 MDIKDRR-HRSLTRGRCGKECRYTSSSLDSEDCRVPAQKSYSSSETLKAYGHDTRMHYGN 59

Qy 59 RVKDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLRYRTDIGLPQCQGYNGAGSDADME 118

Db 60 RVSDLVHRESDEFPRQGTNFTLAELGICEPS-PHRSGYCSDIGILHQGYSLSTGSDADSD 118

Qy 119 ADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHEN----- 161

Db 119 TEGGMSPEHAIRLWGRGIKSSRSSGLSSRENSALTLDSDNENKSDENDEPHTHLSEKLK 178

Qy 162 -----TETDH--- 166

Db 179 DROTSWQQLAETXNSLIRPIPTSSSLLPSAQLPSSHNPVPSQCMPLLDSTNSHQIM 238

Qy 167 -----PGGLQNH---ARLRTP-PPPLSHAHTPNQHHAAASI 197

Db 239 DTNPDEEFSPNSYLLRACSGPQASSSGPSNHHQSQTLRPPLPPPHNHSLS---HHHSSA 295

Qy 198 NSLNRGNFTPRSN---PSPAPTDHLSGEPAGGACBPAAHQENWLLNSNIPLETRNLG 253

Db 296 NSLNENSLNRRNQIHAPAPAPND--LATTP-----ESVQLQDSWVLNSNVPLETR--- 344

Qy 254 KQPFGLTLDNLIELMDILGASRHDGAYSDDGHFLFK-PGGTSPFLCTTSPGYPLTSSTVYS 312

Db 345 -----HFLKTSSTGTTPLFSSSSSPGYPLTSGTVYT 374

Qy 313 PPBRPLPRSTFARPAFNLLKPKSKYCNWKCAALSATVISATLVILLAYFVAMHLFGLNWHL 372

Db 375 PPBRLLPRNTFSRNFALFKLKPCKYCSWKCAALSATAAAVLLAILLAYFIAMHLLGLNWQL 434

Qy 373 QNMEGQMYEI-----TEDTASSWPVPTDVSLYPSGGTGLETDPDRKKGKGTTEGKPSF 424

Db 435 QPADGHTFSNGLRPGAAGABDGAAPPA-----GRG-----PW 467

Qy 425 FPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIVGRKGL 484

Db 468 VTRNSSIDSGETEVRKVTQEVPPGVFWRSQIHIHSQOFLKFNISLGDALFGVYIRRGL 527

Qy 485 PPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSPSSHETGFIQYLDSGIWHLAFY 544

Db 528 PPSHAQYDFMERLDGK-----EKWSVWESPRRRSIQTLVQNEAVFYQYLDVGLMHLAFY 582

Qy 545 NDGKESEVVSFLTATAIESVDNCPSNICYNGDCLISGTHCHFLGLGPDGCRASCPLVCSGN 604

Db 583 NDGKDKEVVSFSTVILDSVDQCPRNCHNGECVSGVCHCFPGFHGADCAKAAACPVLCSGN 642

Qy 605 GQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTICNPGYKGESCEEVDCMD 664

Db 643 GQYSKGTCLCYSGWKGPEDCVPIISQCIDPSCGGHSGCTIEGNCVCSIGYKGENCEEVDCLD 702

Qy 665 PTCGRGVCVRGECCHFCVVGGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPSWTGHDC 724

Db 703 PTCNHHGVCVNGECLCSPGGGINCGLPRAQCPDQCSGHGTYSLDTGLCSDPNWMPDC 762

Qy 725 SIEICAADCGHGVGVGTCRCEDGWMGAACDQACHPRCAEHGTCDGKCECSPGNNGE 784

Db 763 SVEVCSVDCGTHGVCIGGACRCEEGWTGVACDQRVCHPRCTEHGTCKDGKCECREGNGE 822

Qy 785 HCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHCVCCQLGWRGAGCDTSMETACGDSKDN 844

Db 823 HCTIGRQTTGTETDGGCPDLCNNGRCTLQNSWQCVCQGTWRGPGCNVAMETSCADNKN 882

Qy 845 DGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKFLVGR 904

Db 883 EGDGLVDCLVPDCCLOSTQCNLSLLCRGSRDPLDIIQSH--SGSPAVKSFYDRIKLLVCK 940

Qy 905 DSHIIPGENPFDGGHACVIRGOVMTSDGTPLGVNLSFVNNPLFGYTISRQDGSFDLVT 964

Db 941 DSHIIPGENPFSNLSVLRGQVVTDTGTPLGVNVSVFVKYKYGTYTITRODGMFDLVA 1000

Qy 965 NGGISILRRFERAPFITQEHTLWLPWDRFFVFMETIIMRHEENEIIPSCDLSNFARNPVVS 1024

Db 1001 NGSSSLTLHFERAPFMSQERIVMLPWNSFYAMDTLVMKTEENSIPSCDLSGFVRPDPV 1060

Qy 1025 PSPLTSPASSCAEKGPVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLTHFTI 1084

Db 1061 SSPLSTFFSDAPGRNPVPEVQVLEHEIEVPGSSIKLIYLSRRTAGYKSLKIIIMTQSLV 1120

Qy 1085 PFNLMKVHLMVAVEGRLEFRKWFAPAAADLSYFIWDKTVYNQKVFGLSEAFVSVGYES 1144

Db 1121 PLNLTKVHLMVAVEGHLFQKSLASPNLAYTFIWDKTDAYGQKVYGLSDAVSVGFYEYET 1180

Qy 1145 CPDLLLWEKRTTVLQGYEIDASKLGGWSLDKHHALNIQSGLHKGNGENQFVSQQPPVIG 1204

Db 1181 CPSLLIWEKRTALLQGFELDPNLSGWSLDKHHVLNVKSGILHKGNGENQFLTQQPAVIT 1240

Qy 1205 SIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLVYVGFNFYIRRIFFPSGNVTN 1264

Db 1241 SIMNGRRRSISCPSCNGLAEGNKLAPVALAVIGIDGSLFVGDFNYIRRIFFPSRNVTS 1300

Qy 1265 ELRNKDFRHSHPAKYIYLATDPMGSAVFLSDSNSRRRVFKIKSTVVVKDLVKNSEVVAGT 1324

Db 1301 ELRNKEFKHSNNPAHKYIYLAVDPVSGSLYVSDTNSRRRIYKVKSLTGTKDLAGNSEVV 1360

QY	1325	GDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIISTLLGS	1384
Db	1361	GEQCLPFDEARCGDGGKAVDATLMSPRGIAVDKYGLMYFVDATMIRKVDQNGIISTLLGS	1420
QY	1385	NDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDDNNVVLQISENHQVRIVAGR	1444
Db	1421	NDLTAVRPLSCDSSMDVSQVRLEWPTDLAVDPMDNSLYVLENNVILRITENHQVSIAGR	1480
QY	1445	PMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVTTSGEIS	1504
Db	1481	PMHCQVPGID-YSLSKLAIHSALESASAIASHTGVLYISETDEKKINRLRQVTTNGEIC	1539
QY	1505	LVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRIRFIR	1564
Db	1540	LLAGAASDCDCCKNDVNCYSGDDGYATDAILNSPSSLAVAPDGTIYIADLGNIRIRAVS	1599
QY	1565	KNKPFLLNTQNMVELSSPIDQELYLFDTTGKHLYTQSILPTGDYLYNFTYTGDGDIITLITDN	1624
Db	1600	KNRPILNSFNQYEAASPGQEELYVFNADGIHQYTLSLVTGEYLYNFTYSSDNDVTEWMS	1659
QY	1625	NGNMVNVRDSTGMPLWLVPDQVYVWVTMGTSALKSVTTQGHELAMTYHGNISGLLAT	1684
Db	1660	NGNSLKVRDASGMPRHLLMPDNQIVTLAVGTNGGLKLVSTQTLELGLMTYNGNSGILLAT	1719
QY	1685	KSNENGWTTTFYEYDSFGRLTNVTFPTQGVSSFRSDTSSVHVQVETSSK-DDVTITTNLS	1743
Db	1720	KSDETGWTTFYDYDHEGRLTNVTRPTGVVTSLHREMEKSITIDIENSNRDDDDVTIVTNLS	1779
QY	1744	ASGAFYTLLOQOVNRNSYVIGADGSLRLLLLANGMEVALQTEPHLLAGTVNPTVGKRVNUTLP	1803
Db	1780	SVEASYTVVQOVNRNSYQLCNNGTLRVMYANGMSISFHPSEPHVLAGTVTPTIGRCNISLP	1839
QY	1804	IDNGLNLVEWRQORKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRI	1863
Db	1840	MENGLNSIEWRLRKEQIKGVTVFGRKLRVHGRNLLSIDYDRNIRTEKIYDDHRKFTLRI	1899
QY	1864	LYDAQRPSSLWSPSSRLNGVNVTYSPGGYIAGIQGIMSERMEYDQAGRITSRIFADGKT	1923
Db	1900	IYDQLGRPFLWLPSSGLAANVSVFFNGRLAGLQRGAMSSERTIDKQGRIIISERNFADGKV	1959
QY	1924	WSYTYLEKSMVLLHHSORQYIFEFDXNDRLSSVTMPNVARQTLLETIRSVGYENIYQPPE	1983
Db	1960	WSYTYLEKSMVLLLSQORQYIFEYDSSDRLHAVTMPSVARHSMSTHTSVGYIRNIYNPPE	2019
QY	1984	GNASVIOQFTEDEGHLHHTFYLGTGRRVYIKYKLSKLAETLYDTTKVSFTYDETFAGMLKT	2043
Db	2020	SNASVIFYSDDGRILKTSFLGTGRQVFKYKGLSKLSEIYVDSTAVTFGYDETTGVLM	2079
QY	2044	INLQNEGFTCTIRYRQIGPLIDRQIFRFTTEEGMVNARFDYNY-DNSFRVTSMOAVINETP	2102
Db	2080	VNLQSGGFCTIRYRKIGPLVDKQIYRFSEEGMVNARFDYTYHDNSFRIASIKPIISETP	2139
QY	2103	LPIDLRYDDVSGKTEOPGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVOYEIFRSL	2162
Db	2140	LPVDLYRYDEISGKVEHFKFGVIYYDINQIITTAVMTLKHFDTHGRIKEVOYEMFRSL	2199
QY	2163	MYMMTVQYDNMGRVVKKELKVGYPYANTTRYSEYEDADGQLQTVSINDKPLWRYSYDLNGN	2222
Db	2200	MYMMTVQYDSMGRVTRKELKLGYPYANTTKYTYDYDGDGQLQSAVNDRPTWRYSYDLNGN	2259
QY	2223	LHLLSPGNSARLTPLYDIRDIRTRLGDVOYKXMBEDGFLRQRGDIFEYNSAGULLIKAYN	2282
Db	2260	LHLLNPGNSVRLMPLYDLRDIRTRLGDIPIYKIDDDGFLCQRGSDVFEYNSKGLLTRYN	2319
QY	2283	RAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQ	2342
Db	2320	KANGWNVQYRYDGLGRRASCKTNLGHHLQYFYADLHNPTRVTHVYHNSNSEITSLYYDLQ	2379
QY	2343	GHLFAMELSSGDEFYIACDNIQTPLAVFSGTGLMTKQILYTAIGEIMYMDTNENFQIIGY	2402
Db	2380	GHLFAMESSSGEYYIVASDNTGTPLAVFSINGLMTKQLQYTAIGEIIYYDSNDDQLVIGF	2439

QY	2403	HGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHSLWKHLSSSNVMPFNLVMEKNNNPISNSQ	2462
		:       :     :       :     :	
Db	2440	HGGLYDPLTKLVHFTQRDYDVLAGRWTSPDYTMWKNI GREPA - PFNLVMEKSNPNPLSNEL	2498
QY	2463	DIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTOEWDNKSILGV	2522
Db	2499	DLKNYVTDVKSWSLVMGFQLSNIIPGFRAKMYFVSPPYEL --- TESQACENGQLITGV	2554
QY	2523	QCEVQKQLKAFVTLERFDQLYGSTITSQCAPKTKK --- FASSGSVFGVGKVPALKDGR	2578
Db	2555	QQTTERHNQAFMALE --- GQVISKRLHASIREKAGHWFATSTPIIGKIMFAVKGR	2608
QY	2579	VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGR	2638
Db	2609	VTTGISSATDDSRKIASVLSNAHYLEKMHYSIEGKDTHYFVKIGSADSLVTLAMTSGR	2668
QY	2639	RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYG --- TTLDEEKARVLELAR	2695
Db	2669	KVLDSGVNVTVSQPTLLINGRTRRFTNIEFQYSTLLINIRYGLTADTTLDEEKARVLDQAR	2728
QY	2696	QRAVRQAWAREQOORLREGEGLRAWTEGEKQQVLSTGRVQYDGFVIVSVEQYPELSDSA	2755
Db	2729	QRALGSAWAKEQQKARDGREGSRVWTDGEKQQLLNTGRVQGYEGYVLPVEQYPELADSS	2788
QY	2756	NNIHFMRQSEMGRR	2769
Db	2789	SNIQFLRQNEMGKR	2802

RESULT 15

```

US-10-072-012-489
; Sequence 489, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zertusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02

```

RESULT 15  
US-10-072-012-489



Job time : 97 secs

Db	1780	SVEASYTVVQDVNSYQLCNGTLRVMYANGMSISFHSEPHVLACTVPTIGRCNISLP	1839
Qy	1804	IDNGLNLVEWRQRKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRI	1863
Db	1840	MENGLNSIEWLRKEQIKGVTVFGRKLRVHGRNLLSIDYDRNIRTEKIYDDHRKFTLRI	1899
Qy	1864	LYDQAGRPSLWSPSRLNGVNVTVSPGGYIAGIQGIMSERMEYDQAGRITSRIFADGKT	1923
Db	1900	IYDQLGRPFLWLPSSGLAAVNSYFFNGRLAGLQRGAMSERTDIDKQRIISRMEFADGV	1959
Qy	1924	WSYTYLEKSMVLLHLSQRYIFEFDKNDRLSSVTMPNVARQTLTIRSVGYRNIYQPP	1983
Db	1960	WSYTYLEKSMVLLQSQRQYIFEYDSSDRLHAVTMPSPVARHSMSTHTSVGYIRNIYN	2019
Qy	1984	GNASVIQDFTEDGHLHHTFYLGTGRRVVIYKYGKLSKLAETLYDTTKVSFTYDETAGMLKT	2043
Db	2020	SNASVIFDYSDDGRILKTSFLGTGRQVFYKYGKLSKLSSEIVDYDSTAVTFGYDETTGV	2079
Qy	2044	INLQNEGFTCTIRYRQIGPLIDRQIFRETEEGMVNARFDYNY-DNSFRVTSMQAVINETP	2102
Db	2080	VNLQSGGFSCITIRYRKIGPLVDKQIYRFESEGMVNARFDYTYHDNSFRIASIKPIISETP	2139
Qy	2103	LPIDLRYDDVSGKTEQFGKFGVYYDINQIITAVMTHTKHFDAYGRMKEVQYEIFRSL	2162
Db	2140	LPVDLYRYDEISGKVEHFGKFGVYYDINQIITAVMTLSKHFDTHGRIKEVQYEMFRSL	2199
Qy	2163	MYWMTVQYDNMGRVVKELKVGPIYANTTRYSEYDADGQLQTVTSINDKPLWRYSYDLNGN	2222
Db	2200	MYWMTVQYDSMGRVTKRELKLGPIYANTTRYTYDYDGDGQLQSVAVNDRPTWRYSYDLNGN	2259
Qy	2223	LHLLSPGNSARLTPLRYDIRDITRLGDVQYKMDGDFLRQRGDIFEYNSAGLLIKAYN	2282
Db	2260	LHLLNPGNSVRLMPLRYDLRDIRTLRGDIPYKIDDDGFLCQSGSDVFEYNSKGLLTRAYN	2319
Qy	2283	RAGSWSVRYRVDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQ	2342
Db	2320	KANGWNVQYRYDGLGRRASCKTNLGHHLQFYADLHNPTRVTHVYNHSNSEITSLYYDLQ	2379
Qy	2343	GHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAHYIMDTNPNFQIIGY	2402
Db	2380	GHLFAMESSSGEYVVASDNTGTPLAVFINGLMIKQLQYTAHYIYDSNPDFQLVIGF	2439
Qy	2403	HGGLYDPLTKLVHMGRRDYDLAGRWTSPDHKLWKLSSSNVMPFNLYMFKNNPISNSQ	2462
Db	2440	HGGLYDPLTKLVHFTQRDYDLAGRWTSPDYTMWKNIGREPA-PFNLYMFKSNPLSNEL	2498
Qy	2463	DIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDNKSILGV	2522
Db	2499	DLKNYVTDVKSWLVMFGFQLSNIIPGFPRAKNYFVSPPYEL-----TESQACENGQLITGV	2554
Qy	2523	QCEVQQLKAFVTLERFDQLYGSTITSQQAPKTKK----PASSGSVFGKGVKFKALKDGR	2578
Db	2555	QOTTERHNQAFMALE-----QGVISKRLHASIREKAGHWFATSTPIIGKIMFAVKGR	2608
Qy	2579	VTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPSPSEGDLAAILGLSGR	2638
Db	2609	VTTGISSIAATDDSRKIASVLSNAHYLEKMHYSIEGKOTHYFVKIGSADSDLVTLAMTSGR	2668
Qy	2639	RTLENGVNVTVSINTVLNGRTRRYTDIQLQYGALCLNTRYG---TTLDEEKARVLELAR	2695
Db	2669	KVLDSGVNVTVSQPTLLINGRTRRFTNIEFYSTLLINIRYGLTADTLDEEKARVLDQAR	2728
Qy	2696	QRAVRQAWAREQORLREGEEGLRAWTEGKQQVLSTGRVQVGDGFFVISVEQYPELSDSA	2755
Db	2729	QRALGSAWAKEQOKARDGREGSRVWTDGKQQLLNTGRVQVGYEGYVLPVEQYPELADSS	2788
Qy	2756	NNIHFMRRQSEMGRR	2769
Db	2789	SNIQFLRQNEMGKR	2802

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:12:48 ; Search time 85 Seconds  
(without alignments)  
9204.400 Million cell updates/sec

Title: US-10-029-020-14  
Perfect score: 14887  
Sequence: 1 MDVKERKPYRSLTRRRDAER.....ELSDSANNIHFMRQSEMGRR 2769

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES												
Result No.	Score	Query Match	Length	DB ID	Description							
1	14887	100.0	2769	5	ABG70388	Human	TEN					
2	14748.5	99.1	2758	5	ABG97359	Human	CGD					
3	14406.5	96.8	2794	5	ABE98401	Human	NOV					
4	10221	68.7	2721	5	ABP53587	Human	NOV					
5	10201	68.5	2725	5	ABP53586	Human	NOV					
6	9888.5	66.4	2628	5	ABP53588	Human	NOV					
7	9856	66.2	2613	5	ABP53589	Human	NOV					
8	9536	64.1	2733	4	AAU08680	Human	FCT					
9	9531	64.0	2733	7	ADB32024	Human	FCT					
10	9530.5	64.0	2724	4	AAU08681	Human	FCT					
11	9530.5	64.0	2724	7	ADB32029	Human	FCT					
12	9501	63.8	2725	5	ABG61913	Prostate						
13	9314	62.6	2633	6	ABR58318	BCU0205B						
14	8874	59.6	2551	6	ABR58317	BCU0205A						
15	7296.5	49.0	2136	4	AAM78695	Human	pro					
16	6258.5	42.0	1688	6	ABR58344	Human	pro					
17	4920	33.0	1351	6	AAO29571	Human	Pc0					
18	3979.5	26.7	1045	4	AAAB92858	Human	pro					
19	3806.5	25.6	2515	4	ABB71354	Drosophil						
20	3730	25.1	964	4	AAAB93294	Human	pro					
21	3302	22.2	1015	4	AAW79679	Human	pro					
22	2989	20.1	768	2	AAW44817	Human	gam					
23	2989	20.1	768	2	AAAY06639	Gamma-her						
24	2989	20.1	768	3	AAAY71176	Human	Her					
25	2989	20.1	768	4	AAU09891	Human	her					

26	2351	15.8	625	4	AAB92780	Human	pro
27	1574	10.6	527	2	AAW44818	Human	gam
28	1310	8.8	402	4	ABG04674	Novel	hum
29	1244	8.4	459	7	ADE08090	Novel	pro
30	1205	8.1	746	4	ABG07028	Novel	hum
31	1184.5	8.0	777	4	ABB65598	Drosophil	
32	1112	7.5	495	6	ADA54781	Human	pro
33	1043	7.0	1124	4	ABB59241	Drosophil	
34	1037.5	7.0	587	4	ABB59244	Drosophil	
35	976	6.6	242	4	ABG04673	Novel	hum
36	932	6.3	429	4	AAU00392	Human	sec
37	798	5.4	151	3	AAAB40417	Human	ORF
38	792.5	5.3	381	4	AAAB61141	Human	NOV
39	792.5	5.3	381	4	AAU08679	Human	FCT
40	792.5	5.3	381	6	ABO23243	Human	bre
41	792.5	5.3	381	7	ADB32022	Human	FCT
42	682	4.6	865	4	ABG03234	Novel	hum
43	652	4.4	2110	7	AAO30847	Human	cel
44	650	4.4	2108	6	ABO01379	Human	pro
45	649.5	4.4	2167	5	ABG70361	Novel	hum

ALIGNMENTS

RESULT 1  
ABG70388  
ID ABG70388 standard; protein; 2769 AA.  
XX  
AC ABG70388;  
XX  
DT 05-NOV-2002 (first entry)  
XX  
DE Human TEN-M4-like protein.  
XX  
KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;  
KW cell signal processing; metabolic pathway modulation; metabolic disorder;  
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;  
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;  
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;  
KW memory defect; infertility; congenital heart defect; hair growth;  
KW pigmentation disorder; endocrine disorder; respiratory disease; health;  
KW gastro-intestinal disease; reproductive; neurological disease;  
KW bone marrow transplantation; endocrine disease; allergy; inflammation;  
KW nephrological disorder; urinary system disorder; age-related disorder;  
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;  
KW adipocyte complement-related Clq tumour necrosis factor; out at first;  
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;  
KW type 1a membrane sushi-domain containing domain; butyrophilin;  
KW single nucleotide polymorphism.  
Homo sapiens.

Key	Location/Qualifiers
Misc-difference 28	/note= "Asp substituted by Gly as a result of a single nucleotide polymorphism (SNP)"
Misc-difference 64	/note= "Val substituted by Ala as a result of a single nucleotide polymorphism (SNP)"
Misc-difference 76	/note= "Ala substituted by Thr as a result of a single nucleotide polymorphism (SNP)"
WO200257453-A2.	
25-JUL-2002.	
19-DEC-2001; 2001WO-US050331.	
19-DEC-2000; 2000US-0256704P.	
20-DEC-2000; 2000US-0257314P.	

PR 02-MAY-2001; 2001US-0288153P.  
PR 29-MAY-2001; 2001US-0294075P.  
PR 24-JUL-2001; 2001US-0307506P.  
PR 10-AUG-2001; 2001US-0311590P.  
PR 10-AUG-2001; 2001US-0311613P.  
PR 29-AUG-2001; 2001US-0315617P.  
PR 14-SEP-2001; 2001US-0322358P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
PI Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;  
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;  
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;  
XX  
DR WPI; 2002-590744/63.  
DR N-PSDB; ABS52100.  
XX  
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,  
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious  
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or  
PT cancer.  
XX  
PS Claim 1; Page 53; 318pp; English.  
XX  
CC The present invention relates to new NOVX polypeptides. The invention is  
CC useful for treating or preventing a NOVX-associated disorder such as  
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell  
CC signal processing and metabolic pathway modulation in a subject,  
CC preferably human. The invention is also useful for treating metabolic  
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative  
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune  
CC disorders, hematopoietic disorders and various cancers. The molecules of  
CC the invention are also useful for treating or preventing cirrhosis, heart  
CC pancreatitis, learning and memory defects, infertility, congenital heart  
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,  
CC respiratory disease, gastro-intestinal diseases, reproductive, health,  
CC neurological diseases, bone marrow transplantation, endocrine diseases,  
CC allergy and inflammation, neurological disorders, urinary system  
CC disorders, neuropsychiatric disorders and age-related disorders. The  
CC present amino acid sequence represents a NOVX protein of the invention  
XX  
SQ Sequence 2769 AA;  
  
Query Match 100.0%; Score 14887; DB 5; Length 2769;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MDVKERKPYRSLTRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
DB 1 MDVKERKPYRSLTRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
  
QY 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDGLPQCGYSMGAGSDADMEAD 120  
DB 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDGLPQCGYSMGAGSDADMEAD 120  
  
QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLTPP 180  
DB 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLTPP 180  
  
QY 181 PPLSHAHTPNQHHASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQEPAAQENWL 240  
DB 181 PPLSHAHTPNQHHASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQEPAAQENWL 240  
  
QY 241 LNSNIPLETNRLGKQPFGLTQDNLIEMDILGASRDGAYS DGHFLFKPGGTSPLFCTTS 300  
DB 241 LNSNIPLETNRLGKQPFGLTQDNLIEMDILGASRDGAYS DGHFLFKPGGTSPLFCTTS 300  
  
QY 301 PGYPLTSSTVYSPPPRPLPRSTFARPAFNLLKPKSKYCNWKAALSAIVISATLVILLAYF 360  
DB 301 PGYPLTSSTVYSPPPRPLPRSTFARPAFNLLKPKSKYCNWKAALSAIVISATLVILLAYF 360  
  
QY 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLTTPDRKKGKGTTEGK 420  
DB 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLTTPDRKKGKGTTEGK 420

DB 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLTTPDRKKGKGTTEGK 420  
QY 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFRSQVFIHDPVHLKFNVSGLKAAALVGIVG 480  
DB 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFRSQVFIHDPVHLKFNVSGLKAAALVGIVG 480  
QY 481 RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDSGIWH 540  
DB 481 RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDSGIWH 540  
QY 541 LAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCISGTCHCFLGFLGPDCCGRASCPVL 600  
DB 541 LAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCISGTCHCFLGFLGPDCCGRASCPVL 600  
QY 601 CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTICNPGYKGCSEEV 660  
DB 601 CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTICNPGYKGCSEEV 660  
QY 661 DCMPTCSGRGVCVRGECCHCFVWGVTNCPETPRATCLDQSGHGTFLPDTGLCSCDPSWT 720  
DB 661 DCMPTCSGRGVCVRGECCHCFVWGVTNCPETPRATCLDQSGHGTFLPDTGLCSCDPSWT 720  
QY 721 GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCTCRDGKCESPG 780  
DB 721 GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCTCRDGKCESPG 780  
QY 781 WNGEHCITIAHYLDVVKVKEGCPGLCNGNCRCTLDLNGWHVCQGLGWRGAGCDTSMETACGD 840  
DB 781 WNGEHCITIAHYLDVVKVKEGCPGLCNGNCRCTLDLNGWHVCQGLGWRGAGCDTSMETACGD 840  
QY 841 SKNDGDGLVDCMDPDCCCLQPLCHINPLCLGSPNPLDIQETQVPVSQQNLHSHFYDRIKF 900  
DB 841 SKNDGDGLVDCMDPDCCCLQPLCHINPLCLGSPNPLDIQETQVPVSQQNLHSHFYDRIKF 900  
QY 901 LVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRQDGSF 960  
DB 901 LVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRQDGSF 960  
QY 961 DLVTNGGSIILRFERAPFITQHTLWLPWDRFFVMTIMRHEENEIPSCDLSNFARPN 1020  
DB 961 DLVTNGGSIILRFERAPFITQHTLWLPWDRFFVMTIMRHEENEIPSCDLSNFARPN 1020  
QY 1021 PVVSPSLTSSFASSCAEKGPVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT 1080  
DB 1021 PVVSPSLTSSFASSCAEKGPVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT 1080  
QY 1081 HPTIPENLMKVHLMVAVEGRLEPRKWFAPADLSYFIWDXTDVYNQKVFGLSEAFVSVGY 1140  
DB 1081 HPTIPENLMKVHLMVAVEGRLEPRKWFAPADLSYFIWDXTDVYNQKVFGLSEAFVSVGY 1140  
QY 1141 EYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHAALNIQSGILHKNGENQFVSQOP 1200  
DB 1141 EYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHAALNIQSGILHKNGENQFVSQOP 1200  
QY 1201 PVIGSIMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFNYIRRIFFPSGNV 1260  
DB 1201 PVIGSIMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFNYIRRIFFPSGNV 1260  
QY 1261 TNILELRNKDFRSHSHSPAHKYIATDPMGAVFLSDNSRRRVFKIKSTVVVVKDLVKNSEV 1320  
DB 1261 TNILELRNKDFRSHSHSPAHKYIATDPMGAVFLSDNSRRRVFKIKSTVVVVKDLVKNSEV 1320  
QY 1321 VAGTGQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIIST 1380  
DB 1321 VAGTGQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIIST 1380  
QY 1381 LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI 1440  
DB 1381 LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI 1440  
QY 1441 VAGRPVHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVTTTS 1500  
DB 1441 VAGRPVHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVTTTS 1500

QY 1501 GEISLVAGAPSGCDKNDANCDGSGDDGVAKDAKLNTPSSLAVCADGELYVADLGNRI 1560  
Db  
QY 1501 GEISLVAGAPSGCDKNDANCDGSGDDGVAKDAKLNTPSSLAVCADGELYVADLGNRI 1560  
Db  
QY 1561 RFIRKNKPFNTQNMVELSSPIDOELYLFDTTGKHLXTQSLPTGDLVYNFTYTGDDITL 1620  
Db  
QY 1561 RFIRKNKPFNTQNMVELSSPIDOELYLFDTTGKHLXTQSLPTGDLVYNFTYTGDDITL 1620  
Db  
QY 1621 ITDNGNMVNVRRDSTGMPLWLVPDGOVYVWVTMGNTSALKSVTTQGHELAMTYHGNSG 1680  
Db  
QY 1621 ITDNGNMVNVRRDSTGMPLWLVPDGOVYVWVTMGNTSALKSVTTQGHELAMTYHGNSG 1680  
Db  
QY 1681 LLATKSNENGWTTTFEYDSFGRLTNVTFPTGOVSSFRSDTDSSVHVQVETSSKDDVTIIT 1740  
Db  
QY 1681 LLATKSNENGWTTTFEYDSFGRLTNVTFPTGOVSSFRSDTDSSVHVQVETSSKDDVTIIT 1740  
Db  
QY 1741 NLSASGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRV 1800  
Db  
QY 1741 NLSASGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRV 1800  
Db  
QY 1801 TLPIDNGLNLVEWRQKEQARGQVTVFGRRLRVHNRNLLSLDFRVTRTEKIYDDHRKFT 1860  
Db  
QY 1801 TLPIDNGLNLVEWRQKEQARGQVTVFGRRLRVHNRNLLSLDFRVTRTEKIYDDHRKFT 1860  
Db  
QY 1861 LRILYDQAGRPSSLWSPSSRLNGVNVITYSPGGYIAGIQGIMSERMEYDQAGRITSRIFAD 1920  
Db  
QY 1861 LRILYDQAGRPSSLWSPSSRLNGVNVITYSPGGYIAGIQGIMSERMEYDQAGRITSRIFAD 1920  
Db  
QY 1921 GKTWSYTYLEKSMVLLHSQRQYIFEFDKNDRLSSVTMPNVARQTLETIRSVGYRNIYQ 1980  
Db  
QY 1921 GKTWSYTYLEKSMVLLHSQRQYIFEFDKNDRLSSVTMPNVARQTLETIRSVGYRNIYQ 1980  
Db  
QY 1981 PPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGM 2040  
Db  
QY 1981 PPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGM 2040  
Db  
QY 2041 LKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGVMNARFDYNDNPRVTSMQAVINE 2100  
Db  
QY 2041 LKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGVMNARFDYNDNPRVTSMQAVINE 2100  
Db  
QY 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITAVMTHTKHFDAYGRMKEVQYEIFR 2160  
Db  
QY 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITAVMTHTKHFDAYGRMKEVQYEIFR 2160  
Db  
QY 2161 SLMYWMTVQYDNMGRVVKKELKVGPYANTTRYSEYDADGQLQTVSINDKPLWRYSYDLN 2220  
Db  
QY 2161 SLMYWMTVQYDNMGRVVKKELKVGPYANTTRYSEYDADGQLQTVSINDKPLWRYSYDLN 2220  
Db  
QY 2221 GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKXNDEGFLRQGGDIFEYNSAGLLIKA 2280  
Db  
QY 2221 GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKXNDEGFLRQGGDIFEYNSAGLLIKA 2280  
Db  
QY 2281 YNRAGSWSVRYRYDGLGRVSSKSSHHLQFFVADLTNPTKVTHLYNHSSSEITSLYYD 2340  
Db  
QY 2281 YNRAGSWSVRYRYDGLGRVSSKSSHHLQFFVADLTNPTKVTHLYNHSSSEITSLYYD 2340  
Db  
QY 2341 LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYGEIYMDTNPNFQIII 2400  
Db  
QY 2341 LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYGEIYMDTNPNFQIII 2400  
Db  
QY 2401 GYHGGLYDPLTKLVHMGRDDYDVLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNPISN 2460  
Db  
QY 2401 GYHGGLYDPLTKLVHMGRDDYDVLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNPISN 2460  
Db  
QY 2461 SQDIKCFMTDVNSWLLTFGQLHNVIIPGYKPKDMDAMEPSYELIHTQMKTOEWDNSKIL 2520  
Db  
QY 2461 SQDIKCFMTDVNSWLLTFGQLHNVIIPGYKPKDMDAMEPSYELIHTQMKTOEWDNSKIL 2520  
Db  
QY 2521 GVQCEVQKQKAFVTLERFDQLYGSTITSCQQAQPKTKKFASSGSVFGKGVKFKALDGRVT 2580  
Db  
QY 2521 GVQCEVQKQKAFVTLERFDQLYGSTITSCQQAQPKTKKFASSGSVFGKGVKFKALDGRVT 2580

QY 2581 TDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGRT 2640  
Db  
QY 2581 TDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGRT 2640  
Db  
QY 2641 LENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQAVR 2700  
Db  
QY 2641 LENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQAVR 2700  
Db  
QY 2701 QAWAREQQRLREGEGLRAWTEGEKQQVLSTGRVQGYDGFVVISVEQYPELSDSANNIHF 2760  
Db  
QY 2701 QAWAREQQRLREGEGLRAWTEGEKQQVLSTGRVQGYDGFVVISVEQYPELSDSANNIHF 2760  
Db  
QY 2761 MEQSEMGR 2769  
Db  
QY 2761 MEQSEMGR 2769  
Db  
RESULT 2  
ABG97359  
ID ABG97359 standard; protein; 2758 AA.  
XX  
AC ABG97359;  
XX  
DT 16-DEC-2002 (first entry)  
XX  
DE Human CGDD10, INCYTE 7488573CD1.  
XX  
KW Human; cell growth; differentiation; death; CGDD; cancer;  
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;  
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;  
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;  
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;  
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;  
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;  
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;  
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;  
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;  
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;  
KW rheumatoid arthritis.  
XX  
OS Homo sapiens.  
XX  
PN WO200272830-A2.  
XX  
PD 19-SEP-2002.  
XX  
PF 08-FEB-2002; 2002WO-US003715.  
XX  
PR 09-FEB-2001; 2001US-0268111P.  
PR 23-FEB-2001; 2001US-0271175P.  
PR 08-MAR-2001; 2001US-0274503P.  
PR 09-MAR-2001; 2001US-0274552P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;  
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;  
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;  
XX  
DR WPI; 2002-723356/78.  
DR N-PSDB; ABS78652.  
XX  
PT New human proteins associated with cell growth, differentiation and  
PT death, useful for diagnosing, treating or preventing autoimmune or  
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,  
PT atherosclerosis or hepatitis.  
XX  
PS Claim 1; Page 155-161; 181pp; English.  
XX  
CC The invention relates to an isolated polypeptide comprising CGDD1-12  
CC (cell growth, differentiation and death), a naturally occurring amino  
CC acid sequence at least 90% identical to CGDD, a biologically active  
CC fragment or an immunogenic fragment. Also included are the

CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide  
CC comprising a promoter sequence operably linked to the CGDD  
CC polynucleotides, a cell transformed with the recombinant polynucleotide,  
CC a transgenic organism comprising the recombinant polynucleotide, an anti-  
CC CGDD antibody, screening for compounds which bind to/modulate or are  
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a  
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,  
CC agonists and antagonists are useful for diagnosing, treating or  
CC preventing disorders associated with aberrant expression of CGDD,  
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,  
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia  
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental  
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),  
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or  
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in  
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,  
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune  
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,  
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,  
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,  
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic  
CC infections. They are also useful in the assessment of the effects of  
CC exogenous compounds on the expression of nucleic acid and amino acid  
CC sequences of proteins associated with CGDD. The present sequence  
CC represents a CGDD protein

XX  
SQ Sequence 2758 AA;

Query Match 99.1%; Score 14748.5; DB 5; Length 2758;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 2751; Conservative 0; Mismatches 7; Indels 11; Gaps 3;

QY 1 MDVKERKPYRLRRRAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
DB 1 MDVKERKPYRLRRRAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
QY 61 KDIVPOEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDGLPQCYSMGAGSDADMEAD 120  
DB 61 KDIVPOEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDGLPQCYSMGAGSDADMEAD 120  
QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQHARLRTTP 180  
DB 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQHARLRTTP 178  
QY 181 PPLSHAHTPNQOHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPPAGGAQEPHAQENWL 240  
DB 179 PPLSHAHTPNQOHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPPAGGAQEPHAQENWL 238  
QY 241 LNSNIPLERNLGKQFLGTLQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS 300  
DB 239 LNSNIPLERNLGKQFLGTLQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS 298  
QY 301 PGYPLTSSIVYSPPPRPLPRSTFAPAFNLKPKSKYCNWKAALSAIVISATLVILLAYF 360  
DB 299 PGYPLTSSIVYSPPPRPLPRSTFAPAFNLKPKSKYCNWKAALSAIVISATLVILLAYF 358  
QY 361 VAMHFLGLNHLQPMEGQMEYEITEDTASSWPVPTDVSLSVSGGTGLETDPDRKGKGTTEGK 420  
DB 359 VGKHLF--NWLQPMEGQMEYEITEDTASSWPVPTDVSLSVSGGTGLETDPDRKGKGTTEGK 416  
QY 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHPVHLKFNVS LGKALVGIY 480  
DB 417 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHPVHLKFNVS LGKALVGIY 476  
QY 481 RKGLPPSHQTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSPSHETGFIQYLDSGIWH 540  
DB 477 RKGLPPSHQTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSPSHETGFIQYLDSGIWH 536  
QY 541 LAFYNDGKSEVVSFLTATAESVDNCPNCGNGDCISGTCFCFLGFLGPCGRASCPVL 600  
DB 537 LAFYNDGKSEVVSFLTATAESVDNCPNCGNGDCISGTCFCFLGFLGPCGRASCPVL 596  
QY 601 CSGNGQYMKGRCLCHSGWKGAECVDPTNQCIDVACSNHGTCTGTGTCICNPGYKGESCEEV 660

DB 597 CSGNGQYMKGRCLCHSGWKGAECVDPTNQCIDVACSNHGTCTGTGTCICNPGYKGESCEEV 656  
QY 661 DCMPTCSCGRGVCVRGEGCHCFVGGTNCETPRATCLDQCQSGHGTFLPDTGLCSCDPSWT 720  
DB 657 DCMPTCSCGRGVCVRGEGCHCFVGGTNCETPRATCLDQCQSGHGTFLPDTGLCSCDPSWT 716  
QY 721 GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSPG 780  
DB 717 GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSPG 776  
QY 781 WNGEHCITIAHYLDVWVKEGCPGLCNGNGRCTLDLNGWHVCQQLGWRGAGCDTSMETACGD 840  
DB 777 WNGEHCITIAHYLDVWVKEGCPGLCNGNGRCTLDLNGWHVCQQLGWRGAGCDTSMETACGD 836  
QY 841 SKDNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKF 900  
DB 837 SKDNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKF 896  
QY 901 LVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTLVGVNISFVNPLFGYTTISRQDGSF 960  
DB 897 LVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTLVGVNISFVNPLFGYTTISRQDGSF 956  
QY 961 DLVTNGGISIIILRFERAPFITQEHTLWLPWDRFFVMEIIMRHEENIIPSCDLSNFARP 1020  
DB 957 DLVTNGGISIIILRFERAPFITQEHTLWLPWDRFFVMEIIMRHEENIIPSCDLSNFARP 1016  
QY 1021 PVVSPSLTSPASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT 1080  
DB 1017 PVVSPSLTSPASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT 1076  
QY 1081 HPTIPFNLKMHLMVAVEGRLEFRKWFAPAAAPDLSYYFIWKTDDVYNQKVFGLSEAFVSQY 1140  
DB 1077 HPTIPFNLKMHLMVAVEGRLEFRKWFAPAAAPDLSYYFIWKTDDVYNQKVFGLSEAFVSQY 1136  
QY 1141 EYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKXHALNIQSGILHKNGENQFVSQQP 1200  
DB 1137 EYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKXHALNIQSGILHKNGENQFVSQQP 1196  
QY 1201 PVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFSGNV 1260  
DB 1197 PVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFSGNV 1256  
QY 1261 TNILELRNKDFRHSHPAHKYIYATDPMMSGAVFLSDSNRRVFKIKSTVVVKDLVKNSEV 1320  
DB 1257 TNILEL-----SHSPAHKYIYATDPMMSGAVFLSDSNRRVFKIKSTVVVKDLVKNSEV 1309  
QY 1321 VAGTGDCQLPDDTRCGDGGKATEATLNPRGITVDKFLIYFVDGTMIRRIDQNGIIST 1380  
DB 1310 VAGTGDCQLPDDTRCGDGGKATEATLNPRGITVDKFLIYFVDGTMIRRIDQNGIIST 1369  
QY 1381 LLGSNDLTSARPLSCDSVMDISQVRLWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI 1440  
DB 1370 LLGSNDLTSARPLSCDSVMDISQVRLWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI 1429  
QY 1441 VAGRPMHCQVPGIDHFLLSKVATHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS 1500  
DB 1430 VAGRPMHCQVPGIDHFLLSKVATHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS 1489  
QY 1501 GEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSSLAVCAGDGLYVADLGNIRI 1560  
DB 1490 GEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSSLAVCAGDGLYVADLGNIRI 1549  
QY 1561 RFIRKNKPFNLNTQNMVYELSSPIDQELYLFDTTGKHLTYQSLPTGDLYNFTYTGDDITL 1620  
DB 1550 RFIRKNKPFNLNTQNMVYELSSPIDQELYLFDTTGKHLTYQSLPTGDLYNFTYTGDDITL 1609  
QY 1621 ITDNNGMVNVRRDSTGMPLMLVVPDGOVYVWVTMTGNSALKSVTTQGHELAMTYHNSG 1680  
DB 1610 ITDNNGMVNVRRDSTGMPLMLVVPDGOVYVWVTMTGNSALKSVTTQGHELAMTYHNSG 1669  
QY 1681 LLATKSNENGWTTTFYEYDSFGPLTNVTFPTGOVSSFRSDTSSVHVQVETSSKDDVTITT 1740

Db	1670	LLATKSNENGWTFYEYDSFGRLTNVTFGTQVSSFRSDTSSVHVQVETSSKDDVTITT	1729
QY	1741	NLASGAFYTLLOQOVNRNYYIGADGSLRULLLANGMEVALQTEPHLLAGTVNPTVGKRVN	1800
Db	1730	NLASGAFYTLLOQOVNRNYYIGADGSLRULLLANGMEVALQTEPHLLAGTVNPTVGKRVN	1789
QY	1801	TLPIDNGLNLVEWRQKEQARGQVTVFGRRLRVHNRNLLSLDFRVRTTEKIYDDHRKFT	1860
Db	1790	TLPIDNGLNLVEWRQKEQARGQVTVFGRRLRVHNRNLLSLDFRVRTTEKIYDDHRKFT	1849
QY	1861	LRILYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFAD	1920
Db	1850	LRILYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFAD	1909
QY	1921	GKTWSYTYLEKSMVLLLSQRQYIFEFDKNDRLSSVTMPNVARQTLTIRSVGYRNIYQ	1980
Db	1910	GKTWSYTYLEKSMVLLLSQRQYIFEFDKNDRLSSVTMPNVARQTLTIRSVGYRNIYQ	1969
QY	1981	PPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGM	2040
Db	1970	PPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGM	2029
QY	2041	LKTINLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVNARFDYNDNSFRVTSMQAVINE	2100
Db	2030	LKTINLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVNARFDYNDNSFRVTSMQAVINE	2089
QY	2101	TPLPIDLYRYDDVSGKTEQFGKFGVIYYIDINQIITAVMTHTKHFDAYGRMKEVQYEIFR	2160
Db	2090	TPLPIDLYRYDDVSGKTEQFGKFGVIYYIDINQIITAVMTHTKHFDAYGRMKEVQYEIFR	2149
QY	2161	SLMYWMTVQYDNMGRVVKKELKVGYPYANTTRYSYEYDADGQLOQTVSINDKPLWRYSYDLN	2220
Db	2150	SLMYWMTVQYDNMGRVVKKELKVGYPYANTTRYSYEYDADGQLOQTVSINDKPLWRYSYDLN	2209
QY	2221	GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMWDEDGFLRQGGDIFEYNSAGLLIKA	2280
Db	2210	GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMWDEDGFLRQGGDIFEYNSAGLLIKA	2269
QY	2281	YNRAGWSVRYRYDGLCRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSEITSLYYD	2340
Db	2270	YNRAGWSVRYRYDGLCRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSEITSLYYD	2329
QY	2341	LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYGEIYMDTNPNFQIII	2400
Db	2330	LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYGEIYMDTNPNFQIII	2389
QY	2401	GYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNPISN	2460
Db	2390	GYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNPISN	2449
QY	2461	SQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDNKSIL	2520
Db	2450	SQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDNKSIL	2509
QY	2521	GVOCEVQQLKAFVTLERFDQLYGSTITSCQQAPKTKKFASSGSVFGKGVKFAKDGGRVT	2580
Db	2510	GVOCEVQQLKAFVTLERFDQLYGSTITSCQQAPKTKKFASSGSVFGKGVKFAKDGGRVT	2569
QY	2581	TDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDPHYFVKPGPSEGLAILGLSGGRRT	2640
Db	2570	TDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDPHYFVKPGPSEGLAILGLSGGRRT	2629
QY	2641	LENGVNTVTSQINTVLNGRTRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQAVR	2700
Db	2630	LENGVNTVTSQINTVLNGRTRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQAVR	2689
QY	2701	QAWAREQORLREGEGLRAWTEGEKQOVLSTGRVQYDGFVFISVEQYPELSDSANNIHF	2760
Db	2690	QAWAREQORLREGEGLRAWTEGEKQOVLSTGRVQYDGFVFISVEQYPELSDSANNIHF	2749
QY	2761	MRQSEMGRR	2769
Db	2750	MRQSEMGRR	2758

RESULT 3	
ABB98401	
ID	ABB98401 standard; protein; 2794 AA.
XX	
AC	ABB98401;
XX	
DT	21-OCT-2002 (first entry)
XX	
DE	Human NOV1, a TEN-M4 like protein.
XX	
KW	Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW	Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW	Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW	Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW	Gene Therapy; NOV; cancer; heart disease; inflammation;
KW	autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW	asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW	infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW	TEN-M4 like protein.
XX	
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	Misc-difference 379 /note= "Encoded by ATGACGGATT"
XX	
PN	WO200255704-A2.
XX	
PD	18-JUL-2002.
XX	
PF	09-JAN-2002; 2002WO-US000554.
XX	
PR	09-JAN-2001; 2001US-0260417P.
PR	10-JAN-2001; 2001US-0260831P.
PR	28-FEB-2001; 2001US-0272338P.
PR	09-MAR-2001; 2001US-0274876P.
PR	18-APR-2001; 2001US-0284704P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI	Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI	Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
PI	Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI	Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI	Gunther E, Stone DJ;
XX	
DR	WPI; 2002-590674/63.
DR	N-PSDB; ABN85378.
XX	
PT	NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT	treating NOVX-associated disorders e.g. cancer, inflammation, or
PT	Alzheimer's disease, and in chromosome mapping, tissue typing or
PT	pharmacogenomics.
XX	
PS	Claim 1; Page 10-11; 358pp; English.
XX	
CC	The present sequence is the protein sequence for a NOV protein. The NOV
CC	proteins and coding sequences are useful for treating or preventing NOV-
CC	associated disorders or in the manufacture of a medicament for treating
CC	the disorders, such as cancer, heart disease, inflammation, autoimmune
CC	disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC	IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC	(e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC	and other wasting disorders associated with chronic diseases. NOV1 is a
CC	TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX	
SQ	Sequence 2794 AA;

Query Match 96.8%; Score 14406.5; DB 5; Length 2794;  
Best Local Similarity 96.7%; Pred. No. 0;

Matches 2704; Conservative 21; Mismatches 41; Indels 31; Gaps 11;									
QY	1	MDVKERKPYRSLRRRDAERRY	TSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60					
Db	1	MDVKERKPYRSLRRRDAERRY	TSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60					
QY	61	KDIVPQEAEEFCRTGANFTL	RELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD	120					
Db	61	KDIVPQEAEEFCRTGANFTL	RELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD	120					
QY	121	TVLSPHPVRLWGRSTRSGRSS	CLSSRANSNLTLTDTHEHENTETDHPGGLQNHARLRTPP	180					
Db	121	TVLSPHPVRLWGRSTRSGRSS	CLSSRANSNLTLTDTHEHENTETDHPGGLQNHARLRTPP	180					
QY	181	PPLSHAHTPNQHHAAASIN	SLNRGNFTPRSNPSPAPTDHSLSGEBPPAGGAQEPAAHAQENWL	240					
Db	181	PPLSHAHTPNQHHAAASIN	SLNRGNFTPRSNPSPAPTDHSLSGEBPPAGGAQEPAAHAQENWL	240					
QY	241	LNSNIPLERNLGKQPFLGT	LQDNLIEMDILGASRHDGAYS	DGHFLFKPGGTSPLFCTTS	300				
Db	241	LNSNIPLERNLGKQPFLGT	LQDNLIEMDILGASRHDGAYS	DGHFLFKPGGTSPLFCTTS	300				
QY	301	PGYPLTSTVYSPPPRPLPR	STFARPAFNLKPKSKYCNWKCAALS	SAIVISATLVILLAYF	360				
Db	301	PGYPLTSTVYSPPPRPLPR	STFARPAFNLKPKSKYCNWKCAALS	SAIVISATLVILLAYF	360				
QY	361	VAMHLFGLNHLQPMEGOM	YEITEDPASSWPVPTDVS	LYPSGTTGLETDRKKGKGTGK	420				
Db	361	VAMHLFGLNHLQPMEGOM	YEITEDPASSWPVPTDVS	LYPSGTTGLETDRKKGKGTGK	420				
QY	421	PSSFFPEDSFIDSGEIDV	GRRASQKIPPGTFWRSQV	FIDHPVHLKFNVSLGKAALVGIY	480				
Db	421	PSSFFPEASFIDSGEIDV	GRRASQKIPPGTFWRSQV	FIDHPVHLKFNVSLGKAALVGIY	480				
QY	481	RKGLPPSHTQDFVELLDG	RRLLTQEARSLGTPRQSR	GTVPSSSHETGFIQYLDSGIWH	540				
Db	481	RKGLPPSHTQDFVELLDG	RRLLTQEARSLGTPRQSR	GTVPSSSHETGFIQYLDSGIWH	540				
QY	541	LAFYNDGKSESEVVSFLT	TAIESVDNCPNSCYNGDCI	SGTCHCFGLGFLGPCGRASCPVL	600				
Db	541	LAFYNDGKSESEVVSFLT	TAI--LDSW-ALCLGDGECVSGTCHCFGLGPCDSRAACPVL	597					
QY	601	CSGNGQYMKGRCLCHSG	WKGAECDPVTNQCIDVACS	NHGTCTITGTCICNP	PGYKGESCEEV	660			
Db	598	CSGNGQYSGRCLCFSG	WKGTEDVPTQCIDPQCGGRGICIMGSCACNS	SGYKGESCEE	657				
QY	661	DCMDPTCSGRGVCVRGE	CHCFVGGTNCETPRATCL	DQCSGHGTFPLPDTGLCSCDPSWT	720				
Db	658	DCIDPGCSNHGVCIHGE	CHCSPGWGGSNCEILKTMCPQCSGHGTYLQESGSCTCDPNWT	717					
QY	721	GHDCSIEICAADCGGHG	VCVGGTCRCEDGWMGAACD	QARACHPRCAEHGT	CRDGKCECSPG	780			
Db	718	GPDCSNEICSVDCSHG	VCVGGTCRCEEGWTGPACN	QARACHPRCAEHGT	CRDGKCECSPG	777			
QY	781	WNGEHTI--AHYLDRVVK	--EGCPGLCNGNRCRTL	DLNGWHVCVQLGWRGAGCDTSMET	836				
Db	778	WNGEHTISLAHYLDRVV	KLSEGCPLCNGNRCRTL	DLNGWHVCVQLGWRGAGCDTSMET	837				
QY	837	ACGDSKNDGDGLVDCMD	PDCCILQPLCHINPLCLGSPNPLDII	QETQVPVSQQNLHSFYD	896				
Db	838	ACGDSKNDGDGLVDCMD	PDCCILQPLCHINPLCLGSPNPLDII	QETQVPVSQQNLHSFYD	897				
QY	897	RIKFLVGRDSTHIIIPGE	NPFDDGHGACVIRGQVMTSDG	PPLVGWNISFVNNPLFGYTISRQ	956				
Db	898	RIKFLVGRDSTHIIIPGE	NPFDDGHGACVIRGQVMTSDG	PPLVGWNISFVNNPLFGYTISRQ	957				
QY	957	DGSFDLVTNGGSIILR	FERAPPITQEHTLWLPDRFFV	VMETIIMRHEENETPSCDLSNF	1016				
Db	958	DGSFDLVTNGGSIILR	FERAPPITQEHTLWLPDRFFV	VMETIIMRHEENETPSCDLSNF	1017				
QY	1017	ARNPVPVSPSLTSFAS	SCEAKGPIVPEIQALQEEIS	ISGCKMRLSYLSSRTPGYKSVLR	1076				
Db	1018	ARNPVPVSPSLTSFAS	SCEAKGPIVPEIQALQEEIS	ISGCKMRLSYLSSRTPGYKSVLR	1077				

QY	1077	ISLTHPTIPFNLMKVHLM	VAVEGRLFRKWF	AAAAAPDLSYYFIW	KTDVYNQKVFLSEAFV	1136			
Db	1078	ISLTHPTIPFNLMKVHLM	VAVEGRLFRKWF	AAAAAPDLSYYFIW	KTDVYNQKVFLSEAFV	1137			
QY	1137	SVGYEYESCPDLILWEK	RTTVLQGEIDASKLGW	SLDKHHALNIQS-GIL	HKGNGENQF	1195			
Db	1138	SVGYEYESCPDLILWEK	RTTVLQGEIDASKLGW	SLDKHHALNIQS-GIL	HKGNGENQF	1197			
QY	1196	VSQOPPVI	SGIMGNRRRSISCPSC	NGLADGNKLLAPVAL	TCSDGSLYVGDFNYIRRI	1255			
Db	1198	VSQOPPVI	SGIMGNRRRSISCPSC	NGLADGNKLLAPVAL	TCSDGSLYVGDFNYIRRI	1257			
QY	1256	PSGNVTNILEL--RNK	DFRSHSHSPA	KYYLATDPM	SGAVFLSDSNSRRVFKIKSTVVVKD	1313			
Db	1258	PSGNVTNILELVRNK	DFRSHSHSPA	KYYLATDPM	SGAVFLSDSNSRRVFKIKSTVVVKD	1317			
QY	1314	LVKNSEVVAGTGDQCL	PFDDTRCGDGGKATE	ATLTNPR---	GITVDKFGLIYFVDGTMIR	1370			
Db	1318	LVKNSEVVAGTGDQCL	PFDDTRCGDGGKATE	ATLTNPRGPGITVDK	FGLIYFVDGTMIR	1377			
QY	1371	RIDQNGIISTLLGSNDL	TSARPLSCDSVMDISQVR	---LEWPTDLAIN	MDNSLYVLDNN	1427			
Db	1378	RIDQNGIISTLLGSNDL	TSARPLSCDSVMDISQVR	QVHLEWPTDLAIN	MDNSLYVLDNN	1437			
QY	1428	VVLQISENHQVRI	VAGRPMHC	VPIDHFLLSKVAI	HATLESATALAVSHNGVLYIAETD	1487			
Db	1438	VVLQISENHQVRI	VAGRPMHC	VPIDHFLLSKVAI	HATLESATALAVSHNGVLYIAETD	1497			
QY	1488	EKKINRIQV	TTSGEISLVAGAPSG	CDCKNDANCD	CFSGDDGYAKDAKLNTPSSLAVCAD	1547			
Db	1498	EKKINRIQV	TTSGEISLVAGAPSG	CDCKNDANCD	CFSGDDGYAKDAKLNTPSSLAVCAD	1557			
QY	1548	GELYADLGNIRIR	FIRKKNKPF	LNTQNM	YELSSPIDQELYLFDTTGKHL	YQSLEPTGYL	1607		
Db	1558	GELYADLGNIRIR	FIRKKNKPF	LNTQNM	YELSSPIDQELYLFDTTGKHL	YQSLEPTGYL	1617		
QY	1608	YNFTYTGDDITLIT	DNNGNMNVRRD	STGMPLWL	VVPDQVYVVTMTGNS	ALKSVTTQG	1667		
Db	1618	YNFTYTGDDITLIT	DNNGNMNVRRD	STGMPLWL	VVPDQVYVVTMTGNS	ALKSVTTQG	1677		
QY	1668	HELAMTYHGNSGLLA	TAKSNENGW	TFYEYDSF	GRLTNTVPTQGVSSFRSD	TDSSVHVQ	1727		
Db	1678	HELAMTYHGNSGLLA	TAKSNENGW	TFYEYDSF	GRLTNTVPTQGVSSFRSD	TDSSVHVQ	1737		
QY	1728	VETSSKDDVTIT	TNLSASGA	FYTLQDQVR	NSYYIGADGSLRLLLANG	MEVALQTEPHLL	1787		
Db	1738	VETSSKDDVTIT	TNLSASGA	FYTLQDQVR	NSYYIGADGSLRLLLANG	MEVALQTEPHLL	1797		
QY	1788	AGTVNPTVGKRV	NVTLPIDNGLN	LVWRQKEQ	ARGQVTVFGRRLR--	VHNRNLLSLDPD	1844		
Db	1798	AGTVNPTVGKRV	NVTLPIDNGLN	LVWRQKEQ	ARGQVTVFGRRLRVLQ	VHNRNLLSLDPD	1857		
QY	1845	RVTRTEKIYDDHRK	FTLRILYDQ	AGRPSLWS	PSSRLNGVNV	YSPGGYIAGIQRGIMSER	1904		
Db	1858	RVTRTEKIYDDHRK	FTLRILYDQ	AGRPSLWS	PSSRLNGVNV	YSPGGYIAGIQRGIMSER	1917		
QY	1905	MEYDQAGRITS	IRIFADGKT	WSYTYLEK---	SMVLLHSQRQYIF	EFDKNDRLS	SVTMPN	1960	
Db	1918	MEYDQAGRITS	IRIFADGKT	WSYTYLEK	QVQSMVLLHSQRQYIF	EFDKNDRLS	SVTMPN	1977	
QY	1961	VARQLETIRSV	GYRN	IYQPE	GNASVIQDFTED	GHLHTFYLGT	GRRVYKYGKLSKL	2020	
Db	1978	VARQLETIRSV	GYRN	IYQPE	GNASVIQDFTED	GHLHTFYLGT	GRRVYKYGKLSKL	2037	
QY	2021	AETLYD	TTKVSFTY	DET	AGMLKTINLQ	NEGFTCTIRYQIGPL	IDRQIFRTEEGMVNAR	2080	
Db	2038	AETLYD	TTKVSFTY	DET	AGMLKTINLQ	NEGFTCTIRYQIGPL	IDRQIFRTEEGMVNAR	2097	
QY	2081	FDXNDNSFRV	TSMAQ	VINETPL	PIDLRYDD	VSQKTEQ	QKFGVIYYDINQIITAVMT	2140	
Db	2098	FDXNDNSFRV	TSMAQ	VINETPL	PIDLRYDD	VSQKTEQ	QKFGVIYYDINQIITAVMT	2157	





Db 2217 GFLRQRGTETFEYSSKGLLTRVYSKSGWTVIYRYDGLGRRVSSKTSLGQHLQFFYADLT 2276  
QY 2319 NPTKVTHLYNHSSEITSLYDLQHLFAMELSSGDEFYIACNIGTPTPLAVFSGTGLMIK 2378  
Db 2277 YPTRITHVYNHSSEITSLYDLQHLFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLK 2336  
QY 2379 QILYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHLEWKH 2438  
Db 2337 QIQYTAYGEIYFDSNIDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDIEIWK 2396  
QY 2439 LSSSNVMPFNLYMFKNNPNISNSQDIKCFMTDVNSWLLTFFGQLHNVIPGYKXPDMDAME 2498  
Db 2397 I-GKDPAPFNLYMFRNNPNASKIHDVKDYITDVNSWLVTFGFHLHNAIPGFPVKFDLTE 2455  
QY 2499 PSYELIHTQMKTOBWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSCQQAPKTKK 2558  
Db 2456 PSYELV---KSQWDDIPPIFGVQQVARQAKAFSLGKMAEVQ---VSRRRAGGAQS 2507  
QY 2559 ---FASSGSVFGKGVKFKALKDGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVD 2615  
Db 2508 WLWFATVKSLLIGKVMIAVSQGRVQTNVLNFIANEDCIKVAAVLNNAFYLENLHFTIEGKD 2567  
QY 2616 THYFVKPGPSEGLAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCL 2675  
Db 2568 THYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVVNGRTRRFADVEMQFGALAL 2627  
QY 2676 NTRYGTTLDEEKARVLELARQAVRQAWAREQORLREGEEGLRAWTEGKQOVLSTGRVQ 2735  
Db 2628 HVRYGMTLDEEKARILEQARQALARAWAREQORVRDGEGARLWTEGKRLLSAGKVQ 2687  
QY 2736 GYDGFVVISVEQYPELSDSANNIHFMRQSEMGR 2769  
Db 2688 GYDGYVLSVEQYPELADSANNIOFLRQSEIGRR 2721

RESULT 5  
ABP53586  
ID ABP53586 standard; protein; 2725 AA.  
XX  
AC ABP53586;  
XX  
DT 17-DEC-2002 (first entry)  
XX  
DE Human NOV15a protein SEQ ID NO:36.  
XX  
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; chromosome 4.  
XX  
OS Homo sapiens.  
XX  
PN WO200262999-A2.  
XX  
PD 15-AUG-2002.  
XX  
PF 31-DEC-2001; 2001WO-US049976.  
XX  
PR 29-DEC-2000; 2000US-0258928P.  
PR 02-JAN-2001; 2001US-0259415P.  
PR 04-JAN-2001; 2001US-0259785P.  
PR 20-FEB-2001; 2001US-0269814P.  
PR 09-MAR-2001; 2001US-0279863P.  
PR 29-MAR-2001; 2001US-0279832P.  
PR 29-MAR-2001; 2001US-0279833P.

PR 13-APR-2001; 2001US-0283889P.  
PR 18-APR-2001; 2001US-0284447P.  
PR 25-APR-2001; 2001US-0286683P.  
PR 29-MAY-2001; 2001US-0294080P.  
PR 16-AUG-2001; 2001US-0312915P.  
PR 17-AUG-2001; 2001US-0313325P.  
PR 17-SEP-2001; 2001US-0322699P.  
PR 26-NOV-2001; 2001US-0333350P.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;  
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;  
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;  
PI Rastelli L, Casman SJ, Boldog F, Burges CE, Edinger S, Ellerman K;  
PI Gunther E, Smithson G, Millet I, Macdougall JR;  
XX  
DR WPI; 2002-732706/79.  
DR N-PSDB; ABQ82343.  
XX  
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.  
PT  
PS Claim 1; Page 113; 444pp; English.  
XX  
CC The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence represents human NOV15a, which is located on chromosome 4  
XX  
SQ Sequence 2725 AA;  
Query Match 68.5%; Score 10201; DB 5; Length 2725;  
Best Local Similarity 66.9%; Pred. No. 0;  
Matches 1872; Conservative 362; Mismatches 462; Indels 102; Gaps 21;  
QY 1 MDVKERKPYRSLTR-RRDAERYTSSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57  
Db 1 MDVKERRYCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDSSRLLYG 60  
QY 58 SRVKDIVPQEAEEFCRTGANFTLRLGLEEVTPPHGTLRYTDIGLPQCGYSMGAGSDAM 117  
Db 61 NRVKDLVHREADFTROGQNFTRQLGVCEPATRRGLAFCAEMGLPHRGYSISAGSDADT 120  
QY 118 EADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHEN---TETHPGGLQNH 174  
Db 121 ENEAVMSPEHAMRLWGRGVKSGRSSCLSSRSNSALTTLTDTEHKNKSDSENEQPASNOGQS 180  
QY 175 RLRTPPPLSHAHTPNQHHAAISINLRGNFTPRSNPSAPPTDHSLSGEPAGGAQ---- 230  
Db 181 TLQPLPP--SHKQHSQHH-PSITSINRSLNRRNQSPAP-----PAALPAELQTT 229

QY 231 -EPAAHQENWLLNSNIPLETRNLGKQPFGLGTLDQNLNLIEMDILGASRHDGAYSDBGHFLFKP 289  
Db 230 PESVQLQDSWLVGNSVPLESR-----HFLFKT 256  
QY 290 G-GTSPLFCTTSPGYPLTSTVYSPPRPLPRSTFARPAFNLLKPKSKYCNWKCAALSAIV 348  
Db 257 GTGTTPLFSTATPGYTMASGSVYSPTRPLPRNTLSRSFAFKFKSKSKYCNWKCTALCAVG 316  
QY 349 ISATLVILLAYFVAMHLFGLNWHLQPMQGYE---ITEDTASSWVPVPTDVSLYPSGGTG 405  
Db 317 VSVLLAILLSYFIAMHLFGLNWQLQQTENDTFENGKNSDT-----MPTNTVSLPSG--- 368  
QY 406 LETPDRKGKGTTEGKPSFFPEDSFIDSGELDVGRRASQKIPPGTFWRSQVFIHPVHLK 465  
Db 369 -----DNGKLGFTQENTIDSGLDIGNRAIQEIPPGIFWRSOLFIDQPQLK 417  
QY 466 FNVSLGKAAALVGIYGRKGLPPSHT----QPDFVELLDGRRLLTQEARSLEGTPRQSRGTV 521  
Db 418 FNISLQKDALIGVYGRKKLPPSHTQSSPYQDFVELLDGSRLLIAREQSRLLETERAGQAR 477  
QY 522 PPSSHETGFIQYLDSGIWHLAFYNDGKESEVVSFLTTAIESVDNCPSPNCYNGDCLISGTC 581  
Db 478 SVSLHEAGFIQYLDSGIWHLAFYNDGKNAEQVSFNTIVIESVVECPRNCHNGECVSGTC 537  
QY 582 HCFLGFLGPCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTC 641  
Db 538 HCFPGFLGPCSRAACPVLCSGNGQYSGRCLCFSGWKGTEDVPTTQCIDPPQCGRGIC 597  
QY 642 ITGTCICNPYKGESCEEVDMDPTCSGRGVCVRGECHECFVGWGTNCETPRATCLDQCS 701  
Db 598 IMGSCACSSGYKGESCEEADCIDPGCSNHGVCIHGECHCSPGWGGSNCEILKTMCPDQCS 657  
QY 702 GHGTFPLDGLCSDPSWTGHDCSIEICAADCGGHGVCVGGTCRCEDEGWMGAACDORACH 761  
Db 658 GHGTYLQESGCTCDPNWTGPDCSNEICSVDCSGHGVCMGTGTCRCEEGTGPACNQACH 717  
QY 762 PRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVV-----KEGCPGLCNGRCITLDLN 815  
Db 718 PRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDKIVDKIKYKEGCPGLCNSNGRCTLQON 777  
QY 816 GWHCVQCLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPDCCLQPLCHINPLCLGSPNP 875  
Db 778 GGHCVQCPGWRGAGCDVAMETLCTDSKDNEDGLIDCDPDCCLQSSCQNPYCRGLPDP 837  
QY 876 LDIIQETQVPVVSQNLHFSFYDRIKFLVGRDSTHIIPEGNPFDDGGHACVIRGQVMTSDGTP 935  
Db 838 QDIISQSLQSPSQAAKSFYDRISFLIGSDSTHVIPEGSPFNKSLASVIRGQVLTADGTP 897  
QY 936 LVGVNLSFVNPNPLFGYTIISRODGSFDLVNTGGISIIILRPERAPPITQEBTLMPLWDRFFV 995  
Db 898 LIGVNVSFHYPEYGYTITRQDGMFDLVANGGASLTIVFERSPFLTQYHTVWIPWNVFYV 957  
QY 996 METIIMRHEENEIPSCDLSNFARPNPVVSPSLTSPASSCAEKGPVPEIQALQEEISIS 1055  
Db 958 MDTLVNKEENDIPSCDLSGFVRPNPIIVSSPLSTFFRSPSPEDSPIIPETOVLHEETTIP 1017  
QY 1056 GCKMPLSYLSSRTPGYKSVLRISLTHPTIPFNLKVKVLMVAVEGRLFRKWFAPAAADLSYY 1115  
Db 1018 GTDLKLSYLSRRAAGYKSVLKITMTQSIIPFNLKVKVLMVAVVGRFLFQKWFPPASPNNLAYT 1077  
QY 1116 FIWKTIDVYNQVFLSEARFVSUGYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDK 1175  
Db 1078 FIWKTIDAYNQVYGLSEAVSVUGYEYESCLDLTLWEKRTAILQGYELDASNMGGWTLDK 1137  
QY 1176 HHALNIQSGILHKNGENQFVSQQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVAL 1235  
Db 1138 HHVLDVQNGILYKNGENQFISQQPPVSSIMGNRRRSISCPSCNGQADGNKLLAPVAL 1197  
QY 1236 TCGSDGSLYVGDFNYTRIFPSPGNVTNILELRNKDFRSHSPAHKYLATDPMGAVFLS 1295  
Db 1198 ACGIDGSLYVGDFNYVRRIFFPSGNVTSVLELRNKDFRHSNPAHRYLATDPTVDLYVS 1257  
QY 1296 DSNRRRVFKIKSTVVVKDLVKNSEVAVAGTGQCLPFDDTRCGDGGKATEATLTNPRGITV 1355

Db 1258 DTNTRRIYRPKSLTGAKDLTKNAEVVAGTGEOCLPFDEARCGDGGKAVEATLMSPKGMV 1317  
QY 1356 DKFGLIYFVDTMIRRIDQNGIISTLLGSDLTTSARPLSCDSVMDISQVRLEWPTDLAIN 1415  
Db 1318 DKNGLIYFVDTMIRKVDQNGIISTLLGSDLTTSARPLTCDTSMHISQVRLEWPTDLAIN 1377  
QY 1416 PWDNSLYVLDNNVLOISENHQVRIVAGPMPHCQVPGIDHFLLSKVVAIHATLESATALAV 1475  
Db 1378 PWDNSIYVLDNNVLOITENRQVRIAAGPMPHCQVPGVE-YPVGKHAVQTTLESATAIAV 1436  
QY 1476 SHNGVLYIAETDEKKINRIRQVTTSTGEISLVAGAPSGCCCKNDANCDCYQSGDGYAKDAK 1535  
Db 1437 SYSGVLYITETDEKKINRIRQVTTTDEISLVAGIPSECDCKNDANCDCYQSGDGYAKDAK 1496  
QY 1536 LNTPSSLAVCADGELYVADLGNIRIRIRKKNKPFELNTQNMVELSSPIDQELYLFDTTGHK 1595  
Db 1497 LSAPSSLAASPDGTLYIADLGNIRIRAVSKNKPILLNSMNFYEVASPTDQELYIFDINGTH 1556  
QY 1596 LYTQSLPTGDLYNFTYTGDDITLITDNNGNMNVNRRDSTGMPLWLVPDGVVWVTMG 1655  
Db 1557 QYTVSLVTGDLYNFSYNDNDITAVTDSNGNTLRIRRDNPMPVRVWSPDNQVIMLTIG 1616  
QY 1656 TNSALKSVTTQGHHELAMMTYHNSGLLATKSNENGWTTFFEYDSFGRLTNVTFPTGOVSS 1715  
Db 1617 TNGCLKGMTAQGLELVLFTHGNSGLLATKSDETGWTTFDYDSEGRLTNVTFPTGVVTN 1676  
QY 1716 FRSDTDSVHVQVETSSK-DDVTITNLSASGAFYTLLOQDVNSYYIGADGSLRLLLAN 1774  
Db 1677 LHGDMDKAITVDIESSSREEDVSITSNLSSIDSFYTMVQDLRNSYQIGYDGLRIYAS 1736  
QY 1775 GMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVWRQKEQARGQVTVFGRRLRVH 1834  
Db 1737 GLDSHYQTEPHVLVAGTANPTVAKRNMTLPENGQNLVWRFRKEQAQGVKVNFGKLRVN 1796  
QY 1835 NRLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSILWSPSSRLNGVNVITYSPGYIA 1894  
Db 1797 GRNLLSVDFDRTTKTEKIYDDHRKFLRLRIAYDTSGHPTLWLPSSKLMVAVNTYSSTGQIA 1856  
QY 1895 GIQRGIMSERMEYDQAGRITSRIEADGKTWSYTYLEKSMVLLLSHRSORQYIFEFDKNDRLS 1954  
Db 1857 SIQRTTSEKVDYDQGRIVSRVFDGKTWSYTYLEKSMVLLLSHRSORQYIFEYDMWDRLS 1916  
QY 1955 SVTMPNVARQTLTETIRSVGYVYNIYQPEGNASVIQDFTEDGHLHHTFYLGTRRVYKY 2014  
Db 1917 AITMPSVARHTMQTIRSIGYVYNIYNPPESNASIIITDYNEEGLLLQTAFLGTSRRVLFKY 1976  
QY 2015 GKLSKLAETLYDTTKVSFTYDETAGMLKTNLQNEGTCTIRYRQIGPLIDRQIFRPTTE 2074  
Db 1977 RRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLIDRQIFRPS 2036  
QY 2075 GMVNARFDYNDNSFRVISMCAVINETPLPIDLYRVDVSGKTEQFGKFGVYIYDINQII 2134  
Db 2037 GMVNARFDYSYDNSFRVISMCGVINETPLPIDLYQPDIDISGKVEQFGKFGVYIYDINQII 2096  
QY 2135 TTAVMTHTKHFDAYGRMKVEQYEIFRSLMYMTVQYDNMGRVVKKELKVGPYANTTRYSY 2194  
Db 2097 STAVMTYTKHFDAGHRTKEIQEYEIFRSLMYWITIYQYDNMGRVTRKREIKIGPANTTKYAY 2156  
QY 2195 EYDADGQLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDIRLGLDVQYK 2254  
Db 2157 EYDVGQLQTVLYLNEKIMWRYNYDLNGLNHLNPSNSARLTPLRYDLRDIRLGLDVQYR 2216  
QY 2255 MDEDFLRQRGGDIFEYNSAGLLIKAYNRAGSWSVRYDYDGLGRRVSSKSSHHLQFFY 2314  
Db 2217 LDEDFLRQRGTETFEYSSKGLLTRVYSKSGWTVIYRYDGLGRRVSSKSTSLGHLQFFY 2276  
QY 2315 ADLTNPTKVTHLYNHSSSEITSLYDLOQHFLFAMELSSGDEFYIACDNIGTPLAVFSGTG 2374  
Db 2277 ADLTYPTRIITHVYNHSSSEITSLYDLOQHFLFAMEISSGDEFYIASDNTGTPLAVFSSNG 2336  
QY 2375 LMIKQILYTAIGEYIMDTNPNFQIIIGYHGLYDPLTKLVHMGRRDYDVLAGRWTSPDHE 2434

Db	2337	LMLKIQIYTAYGEIYFDSNIDFQLVIGFHGGLYDPLTKLIHGERDYDILAGRWTTPDIE	2396
QY	2435	LWKHLSSNNVMPFNLYMFKNPNPISNQDIKCFMTDVNSWLLTFGFLHNVIPGYKPDPM	2494
Db	2397	IWKRI-GKDPAPFNLYMFRNNNPASKIHDVKDIYITDVNSWLVTFGFLHNAIPGFVPKF	2455
QY	2495	DAMEPSYELIHTQMKTQEWDNKSILGVQCEVQKQKAFVTLLERFDQLYGSTITSQQAP	2554
Db	2456	DLTEPSYELV----KSQWDDPIPIFGVQOQVARQAKAFLSLGKMAEVQ----VSRRRAG	2507
QY	2555	KTKK---FASSGSVFGKGVKALKDGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTI	2611
Db	2508	GAQSWLWFATVKSILGKGVMLAVSQGRVQTNVLNIANEDCIKVAAVLNNAFYLENLHFTI	2567
QY	2612	DGVDTHYFVKPGPSEGDILAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYG	2671
Db	2568	EGKDTHYFIKITTPESDLGTLRLTSGRKALENGINVTVSQSTTVNGRTRRFADVEMQFG	2627
QY	2672	ALCLNTRYGTTLDEEKARVLELARQAVRQAWAREQORLREGEEGLRAWTEGEKQOVLST	2731
Db	2628	ALALHVRVGMTLDEEKARILEQARQARALARAWAREQORVRDGEEGARLWTEGEKROLLSA	2687
QY	2732	GRVQYDGFVIVSQYPELSDSANNIHFMRQSEMGR	2769
Db	2688	GKVQGYDGYVLSVEQYPELADSANNOFLRQSEIGRR	2725
RESULT 6			
ID	ABP53588		
XX	ABP53588	standard; protein; 2628 AA.	
AC	ABP53588;		
XX	17-DEC-2002	(first entry)	
DE	Human NOV15c	protein SEQ ID NO:40.	
KW	Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disease; allergy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; goitre; pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; chromosome 4.		
OS	Homo sapiens.		
PN	WO200262999-A2.		
XX	15-AUG-2002.		
PF	31-DEC-2001; 2001WO-US049976.		
XX	29-DEC-2000; 2000US-0258928P.		
PR	02-JAN-2001; 2001US-0259415P.		
PR	04-JAN-2001; 2001US-0259785P.		
PR	20-FEB-2001; 2001US-0269814P.		
PR	09-MAR-2001; 2001US-0279863P.		
PR	29-MAR-2001; 2001US-0279832P.		
PR	29-MAR-2001; 2001US-0279833P.		
PR	13-APR-2001; 2001US-0283889P.		
PR	18-APR-2001; 2001US-0284447P.		
PR	25-APR-2001; 2001US-0286683P.		
PR	29-MAY-2001; 2001US-0294080P.		
PR	16-AUG-2001; 2001US-0312915P.		
PR	17-AUG-2001; 2001US-0313325P.		
PR	17-SEP-2001; 2001US-0322699P.		
PR	26-NOV-2001; 2001US-0333350P.		
XX			
PA	(CURA-) CURAGEN CORP.		
XX	Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;		
PI	Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;		
PI	Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;		
PI	Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;		
PI	Gunther E, Smithson G, Millet I, Macdougall JR;		
XX	WPI; 2002-732706/79.		
DR	N-PSDB; ABQ82345.		
XX	New NOVX polypeptides and polynucleotides useful for treating NOVX-		
PT	associated disorders, such as cancers, neurological disorders, disorders		
PT	of vesicular transport, gastrointestinal disorders, and autoimmune		
PT	diseases.		
XX	Claim 1; Page 122; 444pp; English.		
CC	The present invention describes novel human proteins designated NOVX,		
CC	where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,		
CC	cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,		
CC	antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,		
CC	immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,		
CC	protozoacide and antihelminthic activities, and can be used in gene		
CC	therapy. The NOVX proteins, nucleotides or antibodies can be used in the		
CC	manufacture of a medicament for treating a syndrome associated with a		
CC	human disease selected from NOVX-associated disorder, such as cancers		
CC	(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,		
CC	ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,		
CC	ischaemic cerebrovascular disease, Alzheimer's disease or Pick's		
CC	disease), disorders of vesicular transport (e.g. cystic fibrosis,		
CC	diabetes mellitus, Grave's disease, or goitre), gastrointestinal		
CC	disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),		
CC	autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic		
CC	anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic		
CC	and protozoal infections. The NOVX proteins can be used as immunogens to		
CC	produce antibodies and as vaccines. The NOVX nucleotide sequences may be		
CC	used in chromosome mapping, identifying individuals from minute		
CC	biological samples (tissue typing), and in forensic identification of a		
CC	biological sample. The present sequence represents human NOV15c, which is		
CC	located on chromosome 4		
XX	Sequence 2628 AA;		
QY	Query Match 66.4%; Score 9888.5; DB 5; Length 2628;		
Db	Best Local Similarity 65.2%; Pred. No. 0;		
QY	Matches 1821; Conservative 348; Mismatches 437; Indels 185; Gaps 21;		
Db	1 MDVKERKPYRSLTR-RRDAERYTSSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57		
QY	1 MDVKERRPYCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDSSRLLYG 60		
Db	58 SRVKDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLRYTDIGLPQCYSYMGAGSDADM 117		
QY	61 NRVKDLVHREADEFTTR-----		
Db	118 EADTVLSPPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLR 177		
QY	77 -----QEQPASNQGST-----LQ----- 90		
Db	178 TPPPPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSGEPAGGAQ-----EP 232		
QY	91 --PLPPSHKHQHSQAHH-PSITSLNRNLTNRNQSPAP-----PAALPAELQTTTPES 139		
Db	233 AHAQENWLLNSNIPLETNRNLGKQPFGLTLDNLIEMDILGASRHDGAYSDDGHFLFKPG-G 291		
QY	140 VQLQDSWVLGNSVPLESR-----HFLFKTGTG 166		
Db	292 TSPLFCTTSPGYPLTSSVYSPPPRPLPRSTFARPAFNLKPKSKYCNWKCAALSAIVISA 351		
QY	167 TPLFSTATPGYTMASGSVYSPPTPLPRNTLSRSFAFKKSSKYCSWKCTALCAVGVSV 226		
Db	352 TLVILLAYFVAMHLFGLNWHLOPMEGQMYE---ITEDTASSWPVPFTDVSLYPSGGTGLT 408		

Db 227 LLAIIISYFIAMHLFGLNWLOQTENDTTFENGKYNSDT-----MPTNTVSLPSG----- 275  
QY 409 PDRKKGTTGKSPSFFPEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIIDHPVHLXFN 468  
Db 276 -----DNGKLGFTQENNTIDSGELDIGRAIQEIPPGIFWRSQIFIDQPQLXFN 327  
QY 469 SLGKAALVGIYGRKGLPPSHTQFDFVELLDGRRLLTQEARSLGTPROSRGTVPSSSHET 528  
Db 328 SLQKDALIGVYGRKGLPPSHTQYDFVELLDGSLRIAREQSLLETERAGRQARSVSLHEA 387  
QY 529 GFQYLDGSIWHLAFYNDGKESEVVSFLTITAEISVDNCPNSCYNGNDCISGTCFCFLGFL 588  
Db 388 GFQYLDGSIWHLAFYNDGKNAEQVSFNTIVIESVVECPRNCHNGECVSGTCHCFPGFL 447  
QY 589 GPCGRASCPVLCNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTITGTCIC 648  
Db 448 GPCDSRAACPVLCSNGQYSGRCLCFSGWKGTEDVPTTQCIDPQCGGRGICIMGSCAC 507  
QY 649 NPGYKGESCEEVDMDPTCSGRGVCVRGEGCHCFVGGTNCETPRATCLDQCSCGHGFLP 708  
Db 508 NSGYKGSCEEADCDPCCSNHGVCIHGECHCSPGWGGSCNCEILKTMCPDQCSCGHGTYLQ 567  
QY 709 DTGLCSCDPSWTGDCSIEICAADCGHGVGVGGTCRCEDGMWGAACDQACHPRCAEHG 768  
Db 568 ESGSCTCDPNWTGPDCSNEICSVDCGSHGVGCMGGTCRCEEGWTGPACNQACHPRCAEHG 627  
QY 769 TCRDGKCECSPGWNGEHCTIAHYLDRVV-----KEGCPGLCNGNGRCTLDLNGWHVCVQ 822  
Db 628 TCKDGKCECSPGWNGEHCTIAHYLDKIVKDKIGYKEGCPGLCNGNGRCTLDQNGGHVCVQ 687  
QY 823 LGWRGAGCDTSMETACGDSKONDGGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIITQET 882  
Db 688 PGWRGAGCDVAMETLCTDSKONEGDLIDCMDPDCCLOSSCQNPYCRGLPDQPDITISQS 747  
QY 883 QVPVSQQNLHSPYDRKFLVGRDSTHILIPGENPFDGGHACVIRGQWMTSDGTPLVGNIS 942  
Db 748 LQSPSQAAKSPYDRISFLIGSDSTHVIPEGSPFNKSLASVIRGQVLTADGTPLIGVNV 807  
QY 943 FVNNPLFGYTISSRQDGSFDLVNTGGISILRFERAPFITQSHTLWLPWDRFFVFMETIMR 1002  
Db 808 FFHYPEGYTITRQDGMFDLVANGGASLTLVFERSPFLTQVHTVWIPWNVFVYVMDLAME 867  
QY 1003 HEENEIPSCDLSNEARNPNVSPSPILTSFASSCAEKGPIVPEIQALQEEISISGCKWRLS 1062  
Db 868 KEENDIPSCDLSGFVRPNPIIVSSPLSTFFRSSPEDSPIPETQVLEHEETIPGTDLKL 927  
QY 1063 YLSSRTPGKSVLRISLTHPTIPENLMKVHLMVAVEGRLPFRKWFAPADLSYFIWDKTD 1122  
Db 928 YLSSRAAGKSVLKITMTQSIIPENLMKVHLMVAVVGRLVQKWFAPASPNNLAYFIWDKTD 987  
QY 1123 VYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTTVLQGEYIDASKLGGWSLDKHHALNIQ 1182  
Db 988 AYNQKVYGLSEAVSVGYEYESCLDLTLWEKRTAILQGEYELDASNMGGWTLDKHHVLDVQ 1047  
QY 1183 SGILHKGNGENQFVSQQPPVIGSINGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGS 1242  
Db 1048 NGILYKNGENQFISQQPPVVSINGNRRRSISCPSCNQADGNKLLAPVALACGIDGS 1107  
QY 1243 LYVGDFNYIRRIFFPSGNVTNILELRNKDFRHSHPAHKYIYLATDPMGAVFLSDSNSRRV 1302  
Db 1108 LYVGDFNYVRRIFPSGNVTSLVLELRNKDFRHSNPAHRYIYLATDPTVTDLYVSDTNTTRI 1167  
QY 1303 FKIKSTVVVKDLVKNSEVVAGTGDQCLPFDDTRCGDGKATEATLTNPRGITVDFKGLIY 1362  
Db 1168 YRPKSLTGAKDLTKNAEVVAGTGEQCLPFDEARCGDGKAVEATLMSPKGMAVDKNGLIY 1227  
QY 1363 FVDGTWIRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLY 1422  
Db 1228 FVDGTWIRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIY 1287  
QY 1423 VLDNNVVLQISENHQVRIVAGRPMCHQVPGIDHFLLSKVAIHATLESATALAVSHNGVLY 1482  
Db 1423 ELV----KSQQWDDIPPIFGVQQQVARQAKAFLSLGKVAEVQ----VSRRRAGGAQSWLW 2417

Db 1288 VLDNNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYSGVLY 1346  
QY 1483 IAEETDEKKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLNTPSSL 1542  
Db 1347 ITETDEKKINRIRQVTTDGEISLVAGIPSECDCKNDANCDCYQSGDGYAKDAKLAPSSL 1406  
QY 1543 AVCADGELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLEFTTGKHLYTQSLP 1602  
Db 1407 AASPDGTLIADLGNIRIRAVSKNKPLLNSMNFYEVASPTDQELYIFDINGTHQYTVSLV 1466  
QY 1603 TGDYLYNFTYTGDDITLITDNNGNMNVVRDSTGMPLWLVPDQVYVWVTMGNTNSALKS 1662  
Db 1467 TGDYLYNFSYNDNDITAVTDSNGNTLIRRDPNRMPVRVVSVDNQVILWLTIGTNGCLKG 1526  
QY 1663 VTTQGHELAMTYHNSGGLLATKSNENGWTTFFEYDYSFGRLTNTVPTGTQVSSFRSDTDS 1722  
Db 1527 MTAQGLELVLFYHNSGGLLATKSDETGWTFFDYDSEGRLTNTVPTGTGVVTLHGDMDK 1586  
QY 1723 SVHVQVETSSK-DDVTITTNLSASGAFYTLLODQVRNSYYIGADGSLRLLLLANGMEVALQ 1781  
Db 1587 AITVDIESSSREEDVSITSNLSSIDSFTVMQDLRNSYQIGYDGSRLRIYASGLDSHYQ 1646  
QY 1782 TEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEWRQKEQARGQVTVFGRRLRVHNRNLLSL 1841  
Db 1647 TEPHVLAGTANPTVAKRNMTLPGENGQNLVEWFRKEQAQGVNFRKLRVNGRNLSSV 1706  
QY 1842 DFRVTRTEKIYDDHRKFTLRILYDQAGRPSPSPSSRLNGVNVVYSPGGYIAGIQRGIM 1901  
Db 1707 DFRDRTTKTEKIYDDHRKFLRLIAYDTSGHPTLWLPSSKLMAVNVYSSTGQIASIQRTT 1766  
QY 1902 SERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLLSHRSORXYIFEFDKNDRLSSVTPNV 1961  
Db 1767 SEKVDYDQGRIVSRVAFADGKTWSYTYLEKSMVLLLSHRSORXYIFEYDMWDRLSAITMPSV 1826  
QY 1962 ARQTLTIRSVGYRNIYQPPENASVIQDFTEDGHLHHTFYLGTRRRVIYKYGKLSKLA 2021  
Db 1827 ARHTMQTIRSIGYRNIYNPPESNASIITDYNEEGLLLOFAPLGTSSRVLFKYRRQTRLS 1886  
QY 2022 ETLYDITKVSFTYDETAGMLKTNLQNEGFTCTIRYRQIQLPLIDRQIFRFTTEGMVNARF 2081  
Db 1887 EILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIQLPLIDRQIFRSEDEGMVNARF 1946  
QY 2082 DYNVDSFRVTSMQAVINETPLPDLRYDDVSGKTEQFGKFGVYIYDINQIITAVMTH 2141  
Db 1947 DYSYDNSFRVTSMQGVINETPLPDLRYQFDDISGKVEQFGKFGVYIYDINQIISTAVMTY 2006  
QY 2142 TKHFDAYGRMKEVOYEIIFRSLMVMWTVQYDNMGRVVKELKVGPIYANTIRYSYEDADGQ 2201  
Db 2007 TKHFDAGRIKEIOYEIIFRSLMYWITIYQYDNMGRVTREIKIGFPANTTKYAYEXDVGQ 2066  
QY 2202 LQTVSINDKPLWRYSYDNLGNLHLLSPGNSARLTPLRDIYDRITRLGDVQYKMDDEDGFL 2261  
Db 2067 LQTVVLEKIMWRYNYDLNGLNHLNPSNSARLTPLRDLDRITRLGDVQYRLDEDDGFL 2126  
QY 2262 RQGGDIFFEYNSAGLLIKAYNAGSWSVRYRYDGLGRRVSSKSSHSHHLOFFFYADLTNPT 2321  
Db 2127 RQGTETIFEYSSKGLLTVYSGSGWTVIYRYDGLGRRVSSKSTSLGQHLQFFFYADLTPT 2186  
QY 2322 KVTHLYNHSSEITSLYYDLQCHLFAMELSSGDEFIACDNIGTPLAVFSGTGLMIKQIL 2381  
Db 2187 RITHVYNHSSSEITSLYYDLQCHLFAMEISSGDEFIASDNTGTPLAVFSSNGLMLKQIQ 2246  
QY 2382 YTAIGEIMYDTPNFQIIGYHGGLYDPLTKLVNMRDDYDLVLAGRWTSDPHELWKLSS 2441  
Db 2247 YTAIGEIFYDSNIDFQLVGFHGGLYDPLTKLHFGERDYDILAGRWTTTPTDIEIWKRI-G 2305  
QY 2442 SNVMPFNLYMFKNNPISNSODIKCFMTDVNSWLLTFGQLHNVI PGYPKPDMDAMEPSY 2501  
Db 2306 KDPAPFNLYMFRNNNFASKIHVDKDIITDVNSWLVTFGFHLHNAIPGFVPKFDLTPEPSY 2365  
QY 2502 ELIHTQMKTOEWDNSKILGVQCEVQKQLKAFVTLERFDQLYGTITSCQAPKTKK--- 2558  
Db 2366 ELV----KSQQWDDIPPIFGVQQQVARQAKAFLSLGKVAEVQ----VSRRRAGGAQSWLW 2417



Db 328 SLQKDALIGVYGRKGLPPSHTOYDFVELLDGSRLLIAPBQRSLLETERAGRQARSVSLHEA 387

Qy 529 GFIOYLDSGIWHLAFYNDGKESEVUSFLTITAESVDNCPNSCYNGDCISGTCHCFGL 588

Db 388 GFIOYLDSGIWHLAFYNDGKNAEQVSFNTIVIESVVECPNCHNGGECVSGTCHCFGL 447

Qy 589 GPDGRASCPVLCSGNGQYMKGRCLCHSGWKGAECVPTNOCIDVACSNHGTCTITGTCIC 648

Db 448 GPDCSRACPVLCSGNGQYSGRCLCFSGWKGTGTECDVPTTQCIDPQCGRGICIMGSCAC 507

Qy 649 NPGYKGESCEEVDMDPTCSGRGVVGRGECHCFVGGTNCETPRATCLDQCSGHGTFPL 708

Db 508 NSGYKGESCEEADCIDPGCSNHGVCIHGECHCSPGWGGSNCEILKTMCPDQCSGHGTYLQ 567

Qy 709 DTGLCSCDPSWTGHDCSIEICAADCGGHGVVCGGTCTCEDGWMGAACDQACHPRCAEHG 768

Db 568 ESGSCTCDPNWTGPDSCNEICSVDCGSHGVCMGTCRCEEWTGPTCNQACHPRCAEHG 627

Qy 769 TCRDGKCECSPGWNGEHCTIAHYLDVVVKEGCPGLCNGNRCRTLDLNGWHVCVQLGWRGA 828

Db 628 TCKDGKCECSHGWNGEHCTI-----EGCPGLCNSNGRCRTLDQNGWHVCVQPGWRGA 678

Qy 829 GCDTSMETACGDSKNDGDGLVDCMDPDCCIQPLCHINPLCLGSPNPLDIIQETQVPVSQ 888

Db 679 GCDVAMETLCTDSKDNEDGLIDCMDPDCCIQSSCQNPYCRGLPDQDIISQSLQSPSQ 738

Qy 889 QNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOVMTSDGTPLVGNISFVNPL 948

Db 739 QAAKSFYDRISFLIGSDSTHIVGESPFNKSLASVIRGOVLTADGTPLIGVNVSFHYPE 798

Qy 949 FGYTISRQDGSFDLVTNGGIIILRPERAPFITQEHTLWLPWDRFFVMETIIMRHEENEL 1008

Db 799 YGYTITRQDGMFDLVANGGASLTLVFERSPFLTQYHTVWIPNVVYVMDTLVMEKEENDI 858

Qy 1009 PSCDLSNFARPNPVVSPSLTSEASSCAEKGPVPIPEIQALQEBISISGCKMRLSYLSRT 1068

Db 859 PSCDLSGFVRNPPIIVSSPLSTFFRSPSPDSPIIPETQVLHEETIPGTDKLSYLSRA 918

Qy 1069 PGYKSVLRISLTHPTIPFNLKMKVHLMVAVEGRLEFRKWFAPADLSYFIWDKTDVYNQKV 1128

Db 919 AGYKSVLKITMTQSIIPFNLKMKVHLMVAVVGRLEFKWFPPASPFLAYTFIWDKTDAYNQKV 978

Qy 1129 FGLSEAFVSQYVEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHHALNIOGILHK 1188

Db 979 YGLSEAVSVGYEYESCLDLTLWEKRTAILQGYELDASNMGGWTLDKHHVLDVQNGILYK 1038

Qy 1189 NGENQFVSQPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGD 1248

Db 1039 NGENQFISQPPVSSIMNGRRRSISCPSCNQADGNKLLAPVALACGIDGSLYVGD 1098

Qy 1249 NYIRRIFFSGNVNILELRNKDPRHSHSPAKEYLATDPMGAVFLSDSNSRRVFKIKST 1308

Db 1099 NYVRRIFPSGNVTSVLELRNKDPRHSSNPAHRYLATDPVTGDLVSDTNRIRYRPSL 1158

Qy 1309 VVVKDLVKNSEVAVAGTGDQCLPDDTRCGDGKATEATLTNPRGITVDKFLYFVDGTM 1368

Db 1159 TGAKDLTKNAEVAVAGTGEQCLPDEARCGDGKAVEATLMSPKGMAVDKNGLYFVDGTM 1218

Qy 1369 IRRIDONGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNV 1428

Db 1219 IRKVDONGIISTLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDNNV 1278

Qy 1429 VLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVATHATLESATALAVSHNGVLYIAETDE 1488

Db 1279 VLQITENRQVRIAAGRPMHCQVPGVE-YPVVGKHAQTLESATAIAVSYGVLITYETDE 1337

Qy 1489 KKNIRIQVTTSGEISLVAGAPSGCDCKNDANCDGSDGYAKDAKLNTPSSLAVCADG 1548

Db 1338 KKNIRIQVTTDGEISLVAGIPSECDCKNDANCDQYSGDGYAKDAKLAPSSLAASPDG 1397

Qy 1549 ELYVADLGNIRIRFIRKNKPFNLNTQNMVELSSPIDQELYLEFTTGXHLYTQSLPTGDYLY 1608

Db 1398 TLYIADLGNIRIRAVSKNPKLLNSMNFYEVASTDQELYIFDINGHQYTVSLVTGDYLY 1457

Qy 1609 NPTYTGDDITLITDNNGNMNVNRDSTGMPLWLVPDQGVYVMTGMTNSALKSVTTQGH 1668

Db 1458 NFSYSDNDNDITAVTDSNGNTLRIRDRDPNRPVVRVSPDNQVILWTITGNTGLKGMTAQGL 1517

Qy 1669 ELAMWYTHGNSGLLATKSNENGWTTFFYEYDSFGLTNVTFTPTQVSSFRSDTSSVHVQV 1728

Db 1518 ELVFTYHGNUGLLATKSDDETGWTTFFDYDSEGRLTNVTFTPTGVVTLNHLGDMDKAITVDI 1577

Qy 1729 ETSSK-DDVTITNLSASGAFYTLQDQVRNSYIIGADGSLRLLLLANGMEVALQTEPHLL 1787

Db 1578 ESSSREEDVSIITSNLSSIDSFYTMVQDQLRNSYQIGYDGLRIIYASGLDSHYQTEPHVL 1637

Qy 1788 AGTVNPTVGRNVTLPIDNLGNLVWRQREKQARGQVTVFGRRLRVHNRNLLSLDFDRVT 1847

Db 1638 ACTANPTVAKRNMTPGNGQNLVWRFRKEQAQGVKVNFGRLKRVNGRNLLSVDFDRTT 1697

Qy 1848 RTEKIYDDHRKFTLRILYDQAGRPSSLWSPSSRLNGVNVITYSPGGYIAGIQRGIMSERMEY 1907

Db 1698 KTEKIYDDHRKFLRIAYDTSGHPTLWLPSSKLMVNVITYSSTGQIASIQRTTSEKVDY 1757

Qy 1908 DQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRSLSVTMPNVARQTLE 1967

Db 1758 DGQGRIVSRVADGKTWSYTYLEKSMVLLHLSQRQYIFEYDMWRLSAITMPSVARHTMQ 1817

Qy 1968 TIRSVGYRNIYOPPEGNASVIQDFTEDGHLHTFYLGTRRVYKYKLSLAETLYDT 2027

Db 1818 TIRSIGYRNIYNPPESNASIITDYNEGLLLQTAFLGTSRRVLFKYRRQTRLSEILYDS 1877

Qy 2028 TKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEEGMVNARFDYNDN 2087

Db 1878 TRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLIDRQIFRSEDMVNARFDYSYDN 1937

Qy 2088 SFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINOIITAVMTHTKHEDA 2147

Db 1938 SFRVTSMQVINETPLPIDLYQFDDISGKVEQFGKFGVIYDINOIISTAVMTYTKHEDA 1997

Qy 2148 YGRMKEVQYEIFRSLMYWMTVOYDNMGRVVKELKVGVPYANTTRYSEYDADGQLQTVSI 2207

Db 1998 HGRKEIQYEIFRSLMYWITIYQDNMGRVTKREIKIGPFANTTKYAYEYDVGQLQTVYL 2057

Qy 2208 NDKPLWRYSYDLNGLNLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFLRQGGD 2267

Db 2058 NEKIMWRYNYDLNGLNLLNPSNSARLTPLRYDLRDRITRLGDVQYRLDEDEGFLRQGGTE 2117

Qy 2268 IFEYNSAGLLIKAYNRAGSVRYRYDGLGRRVSSKSSHSHHLQFFYADLTNPTKVTHLY 2327

Db 2118 IFEYSSKGLLTRVYSGSGWTVIYRYDGLGRRVSSKTSGLQHLOFFYADLTYPTRITHY 2177

Qy 2328 NHSSEITSLYDLQHLFAMELSSGDEFYIACDNIGTFLAVFSGTGLMIKQILYAYGE 2387

Db 2178 NHSSEITSLYDLQHLFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLKQIQYAYGE 2237

Qy 2388 IYMDTNPNFQIIIGYHGLYDPLTKLVHMGRDYDVLAGRWTSPDHLMKHLSSSNVMPF 2447

Db 2238 IYFSDNIDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDIEIWKRI-CKDPAPF 2296

Qy 2448 NLYMFKNNPISNSQDIKCFMTDVNSWLLTFFGQLHNVIPIGYPKPDMDAMEPSYELIHTQ 2507

Db 2297 NLYMFRNNNPASKIHDVKDYITDVNSWLVTFGFLHNAIPGFVPKFDLTPEPSYELV- 2353

Qy 2508 MKTQEWDNKSIILGVQCEVQKQKAFVTLERFDOLYGSTITSCQAPKTKK---FASSGS 2564

Db 2354 -KSQWDDIPPIFGVQVQVARQAKAFLSLGKMAEVQ----VSRRRAGGAGQSWLWFATVKS 2408

Qy 2565 VFGKGVKFAKDKGRVTTDIISVANEDGRRVAAIINHAHYLENLHFTIDGVDTHYFVKPGP 2624

Db 2409 LIGKGVMLAVSQGRVQTNVNIANEDCIKVAALNNAFYLENLHFTIEGKDTHYFIKTTT 2468

Qy 2625 SEGDLAILGLSGGRRLIENGWNVTVSQINTVLNGRTRRYTDIQLQYCALCLNTRYGTTLD 2684

Db 2469 PESDLGTLRLTSGRKALENGINVTVSQSTTVNGRTRRFADVEMQFALALHVRGTM 2528



QY 762 PRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHVCV 821  
DB 731 PRCIEHGTCKDGKCECREGNGEHCTIGRQTAGTETDGCPLCNGNGRCTLGQNSWQCVCV 790  
QY 822 QLGWRGAGCDTSMETACGSKNDGDGLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQE 881  
DB 791 QTGWRGPGCNVAMETSCADNKDNEGDGLVDCPLDCCCLQSACQNSLLCRGSRDPLDIIQ 850  
QY 882 TQVPVVSQONLHSEFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOVMTSDGTPLVGVNI 941  
DB 851 GQ--TDWPAVKSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOVMTSDGTPLVGVNV 908  
QY 942 SFVNNPLFGYTTISROGSPDLVTNGGISIIILRPERAPFITOEHTLWLPWDRFFVMEIIM 1001  
DB 909 SFVKYKPGYTTITRODGTFDLIANGGASLTILHFERAPFMSQERTVWLPWNSFYAMDITLV 968  
QY 1002 RHENEIPSCDLNFPARPWPVSPPLTSFASSCAEKGPVPEIQALQEEISISGCKMRL 1061  
DB 969 KTEENSIPSCDLNFPARPWPVSPPLTSFASSCAEKGPVPEIQALQEEISISGCKMRL 1028  
QY 1062 SYLSSRTPGYKSVLRISLTHPTIPENLMKVHLMVAVEGRLEFRKWFAPAAADLSYFIWDKT 1121  
DB 1029 RYLSSRTAGYKSVLRISLTHPTIPENLMKVHLMVAVEGRLEFRKWFAPAAADLSYFIWDKT 1088  
QY 1122 DVYNQKVFGLSEAFVSVGYEYESPCDLLILWEKRTTVLQGEYDIEASKLGGWSLKKHALNI 1181  
DB 1089 DAYGQRYVGLSDAVVSVGFYEYETCPSLILWEKRTALLQGFELDPNSNLGGWSLKKHALNV 1148  
QY 1182 QSGILHKGNGENQVSPQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDG 1241  
DB 1149 KSGILHKGNGENQVSPQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDG 1208  
QY 1242 SLVVGDFNYIRIPPSGNVTNILELRNDFRHSHPAHKYIYLAIDPMGSAVFLSDSNRR 1301  
DB 1209 SLVVGDFNYIRIPPSGNVTNILELRNDFRHSHPAHKYIYLAIDPMGSAVFLSDSNRR 1268  
QY 1302 VFQKSTVVVKDLVKNSEVAVAGTQDQCLPFDTRCGDGGKATEATLNPRGITVDFKGLI 1361  
DB 1269 IYRVKSLSGTKDLAGNSEVAVAGTQDQCLPFDTRCGDGGKATEATLNPRGITVDFKGLI 1328  
QY 1362 YFVDTGMIRIDQNGIISTLLGNDLTSARPLSCDSVMDISQVRLWPTDLAINPMDNSL 1421  
DB 1329 YFVDTGMIRIDQNGIISTLLGNDLTSARPLSCDSVMDISQVRLWPTDLAINPMDNSL 1388  
QY 1422 YVLDDNNVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVL 1481  
DB 1389 YVLENNVILRIENHQVSIAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVL 1447  
QY 1482 YIAETDEKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDPFGSDGDAYAKDKNTPSS 1541  
DB 1448 YITETDEKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDPFGSDGDAYAKDKNTPSS 1507  
QY 1542 LAVCAGELYVADLGNIRIRFIRKNKPFLLNTQNMVELSSPIDQELYLFDTTGKHLTYQSL 1601  
DB 1508 LAVAPDGTIYIADLGNIRIRAVSKNKPVLNAFNQYEAASPGQELVYFNADGIHQYTVSL 1567  
QY 1602 PTGDYLYNFTYDGDITLITDNNGMVNVRRDSTGMPLWLVPDQVYVWVTMTNSALK 1661  
DB 1568 VTGEYLYNFTYDNDVTELIDNNGNSLKIRRDSSGMPHLLMPDNQIITLTVGTNGGLK 1627  
QY 1662 SVTTQGHELAMMTYHNGSGLLATKSNENGWTTFFEYDSEFGRLTNVTFTPTGOVSSFRSDTD 1721  
DB 1628 VVSTQNLGLMTYDGTGLLATKSDGTGTTFFDYDHEGRLTNVTTRPTGVVTSLHREME 1687  
QY 1722 SSVHVQVETSSK-DDVTITNLSASGAFYTLQDQVRNSYVIGADGSLRLLLLANGMEVAL 1780  
DB 1688 KSIDIDIENSRDDDDVTITNLSVEASYTVVQDQVRNSYVIGADGSLRLLLLANGMEVAL 1747  
QY 1781 QTEPHLLAGTVNPTVGRNVILPIDNGLNLVEWRQREKQARGQVTVFGRRLRVHNRNLLS 1840  
DB 1748 HSEPHVLAGTITPTIGRCNISLPMENGLNSIEWRLRKEQIKGKVTIFGRKLRVHNRNLLS 1807  
QY 1841 LDFDVRTTEKIYDDHRKFTLRIYDQAGRPSLWSPPSRLNGVNVITYSPGGYIAGIQRGI 1900

DB 1808 IDYDRNIRTEKIYDDHRKFTLRIYDQAGRPSLWSPPSRLNGVNVITYSPGGYIAGIQRGI 1867  
QY 1901 MSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLLSHQRYIFEDKNDRLSSVTMPN 1960  
DB 1868 MSERTDIDKQGRIVSRMFADGKWSYSYLDKSMVLLLSHQRYIFEDKNDRLSSVTMPN 1927  
QY 1961 VARQTLTETIRSVGYRNIIYQPPENASVTDQPTEDGHLHLLHTFYLGTGRRVIYKYGLSKL 2020  
DB 1928 VARHSMSTHTSIGYIRNIYNPPENASVTFDYSDDGRILKTSFLGTGRQVYKYGLSKL 1987  
QY 2021 AETLYDTTKVSFTYDETAGMLKTLINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGVMNAR 2080  
DB 1988 SEIVYDSTAVTFGYDETGTGLKMNLSQGGFSCITIRYKIGPLVDKQIYRFSEEGVMNAR 2047  
QY 2081 FDYNY-DNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITAVM 2139  
DB 2048 FDYTHDNSFRIASIKPVISETPLPLVDLYRYDEISGKVEHFGKFGVIYDINQIITAVM 2107  
QY 2140 THTKHFDAYGRMKVEQYEIFRSLMYMVTVOYDNMGRVVKELKVGYPYANTTRYSEYDAD 2199  
DB 2108 TSKHFDTHGRIKEVQYEMFRSLMYMVTVOYDNMGRVVKELKVGYPYANTTRYSEYDAD 2167  
QY 2200 GOLQTVSINDKPLWRYSYDNLGNLHLLSPGNSARLTPLRYDIRDITRLGDVQYKMDG 2259  
DB 2168 GOLQSVAVNDRPTWRYSYDNLGNLHLLSPGNSARLTPLRYDIRDITRLGDVQYKMDG 2227  
QY 2260 FLRQGGDIFEYNSAGLLIKAYNRAGSWRYRYDGLGRRVSSKSSHHLQFFYADLTN 2319  
DB 2228 YLCQSGSDIFEYNSAGLLIKAYNRAGSWRYRYDGLGRRVSSKSSHHLQFFYADLTN 2287  
QY 2320 PTKVTHLYNHSSEITSLYYDLQHLFAMELSSGDEFYIACNIGTPLAVFSGTGLMIKQ 2379  
DB 2288 PTRITHVYNSHNSSEITSLYYDLQHLFAMELSSGDEFYIACNIGTPLAVFSGTGLMIKQ 2347  
QY 2380 ILYTAYGEIYNDTPNFQIIGYHGGLYDPLTKLVHMGRDDYDLVLAGRWTSPDHKLWHL 2439  
DB 2348 LQYTAYGEIYDSDNPDQFQVIGFHGGLYDPLTKLVHFTQRYDVLGRWTSPDYTMWKNV 2407  
QY 2440 SSSNVMPFNLYMFKNNPNISNQDKCFMTDVNSWLLTFGQLHNVIPGYKPKPMDAMEP 2499  
DB 2408 GKEPA-PFNLYMFKNNPNISNQDKCFMTDVNSWLLTFGQLHNVIPGYKPKPMDAMEP 2466  
QY 2500 SYELIHTQMKTOEWDNSKISILVQCEVQKQLKAFVTLERFDQLYGSTITSCQQAQPKTKK- 2558  
DB 2467 PYELSESQAS----ENGQLITGVQOTTERHNOAFMALE-----GOVITKHLHASIREKA 2516  
QY 2559 ---FASGSGVFGKGVKFAKLDGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVD 2615  
DB 2517 GHWFATTTPIIGKIMFAIKRGRVTTGVSIIASEDSRKVASVLNNAYYLDKMHYSIEGKD 2576  
QY 2616 THYFVKPGPSEGDLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCL 2675  
DB 2577 THYFVKPGPSEGDLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCL 2636  
QY 2676 NTRYG---TTLDEEKARVLELAPQAVRQAWAREQQRLREGEGLRAWTEGEKQQVLSTG 2732  
DB 2637 SIRYGLTPDLDLDEEKARVLDQARALGTAWAKEQQKARDGREGSRLWTEGEKQQVLSTG 2696  
QY 2733 RVQYDGFVVISVEQYPELSDSANNIHFMRQSEMGR 2769  
DB 2697 RVQYDGFVVISVEQYPELSDSANNIHFMRQSEMGR 2733

RESULT 9  
ADB32024  
ID ADB32024 standard; protein; 2733 AA.  
XX  
AC ADB32024;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Human FCTR3b polypeptide.

XX Human; FCTR3b; colorectal cancer; adenomatous polyposis coli;  
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;  
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;  
KW renal cell carcinoma; melanoma; clear cell carcinoma;  
KW granular cell carcinoma; neurological disorder;  
KW neurodegenerative disorder; nerve trauma;  
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;  
KW Gardner syndrome; mental health condition; immunological disorder;  
KW allergy; asthma; lung disease; reproductive disorder; deafness;  
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;  
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;  
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;  
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;  
KW lattice corneal dystrophy.  
XX Homo sapiens.  
OS  
XX US2003087816-A1.  
PN  
XX 08-MAY-2003.  
PD  
XX 05-MAR-2001; 2001US-00800198.  
PF  
XX 03-MAR-2000; 2000US-0186592P.  
PR  
XX (VERM/) VERMET C.  
PA (FERN/) FERNANDES E.  
PA (SHIM/) SHIMKETS R.  
PA (HERR/) HERRMANN J.  
PA (MAJU/) MAJUMDER K.  
PA (MACD/) MACDOUGALL J.  
PA (MISH/) MISHRA V.  
PA (MEZE/) MEZES P S.  
PA (RAST/) RASTELLI L.  
XX Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K,  
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;  
XX WPI; 2003-625633/59.  
DR N-PSDB; ADB32023.  
XX New FCTRX polypeptide and encoding polynucleotide, useful for preventing  
PT or treating FCTRX-related disorders, such as cancer, autoimmune,  
PT neurodegenerative, gastrointestinal, reproductive and inflammatory  
PT diseases.  
XX Claim 1; Page 32-33; 155pp; English.  
PS The invention relates to FCTRX polypeptides and the polynucleotides  
XX encoding them. The sequences of the invention are useful for the  
CC manufacture of a medicament for diagnosing and treating disorders  
CC associated with the FCTRX polypeptide, such as colorectal cancer,  
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal  
CC autoimmune thrombocytopenia, malignant ovarian tumours, renal cell  
CC tumours, mammary tumours, human gliomas, astrocytomas, clear cell  
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear  
CC cell and granular cell carcinomas, neurological disorders,  
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic  
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health  
CC conditions, immunological disorders, allergy and infection, asthma, lung  
CC diseases, male and female reproductive disorders, deafness, glycoprotein  
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis  
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni  
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,  
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence  
CC represents an FCTRX polypeptide of the invention.  
XX Sequence 2733 AA;

Query Match 64.0%; Score 9531; DB 7; Length 2733;  
Best Local Similarity 61.4%; Pred. No. 0;  
Matches 1755; Conservative 413; Mismatches 477; Indels 212; Gaps 25;

QY	1	MDVKERKPYRSLTRRR-DAERRYTSADSEEGKAP-QKSYSSSETLKAYDQDARLAYGS	58
Db	1	MDVKDRR-HRSLTRGRCGKXECRYTSSSLDSEDCRVPTQKSYSSSETLKAYDHSRMHYGN	59
QY	59	RVKDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGFLYRTDIGLPQCQGYSMGAGSDADME	118
Db	60	RVTDLIHRESDEFPRQGTNFTLAEIGICEPS-PHRSGYCSDMGILHQYSLSTGSDADSD	118
QY	119	ADTVLSPEHPVRLWGRSTFSGRSSCLSSRANSNLTLTDEHENTETDHPG-----	168
Db	119	TEGMSPEHAIRLWGRGIKSRSSGLSSRENSALTITDSDNENKSDDENGRPIPTSSPS	178
QY	169	-----G	169
Db	179	LLPSAQLPSSHNPFPVSCQPLLDSTSHQIMDTNDEEFPNSYLLRACSGPQOASSSG	238
QY	170	LQNH---ARLRTPPPLSHAHTPNQHHAAASINSLNRCNFTPRSN---PSPAPTDHSLSG	222
Db	239	PPNHHSQSTLRPPLPP-PHNHTLSHHH-SSANSLNRLNRRSQIHAPAPAPND--LAT	294
QY	223	EPPAGGAQEPHAQENWLNLSNIPLETRNLGKQPFGLCTLQDNLIEMDILGASRHDGAYSD	282
Db	295	TP-----ESVQLQDSWVLNSNVPLETR-----	316
QY	283	GHFLFK-PGTSPLFCTTSPGYPLTSSTVYSPPPRPLPRSTFARPAFNLKPKSKYCNWKC	341
Db	317	-HFLFKTSSGSTPLFSSSPGYPLTSGTVYTPPRLPRNTFSRKAFKLKPKSKYCSWKC	375
QY	342	AALSAIVISATLVILLAYFYAMHLFGLNWHLQPMEGQWEITEDTASSWPVPTDVSLYPS	401
Db	376	AALSAIAAALLAILLAYFI-----VPWSL-----	400
QY	402	GGTGLETPDRKGKGTTEGKPSFPEDSFIDSGEIDYRRASQKIPPGTFWRSQVFIIDHP	461
Db	401	-----KNSSIDSGEAEVGRVTVQVPPGVFWRSQIHISQP	435
QY	462	VHLKFNVS LGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTV	521
Db	436	QFLKFNISLGKDALFGVIYIRGLPPSHAQYDFMERLDGK-----EKWSVVESPRRRSIQ	490
QY	522	PPSHETGFIQYLDSGIWHLAFYNDGKSEVVSFLTITATESVDNCPNSCYNGDCISGTC	581
Db	491	TLVQNEAVFYQYLDVGLWHLAFYNDGDKEMVSFNTVLDSDQDCPRNCHGECVSGVC	550
QY	582	HCFLGLPGDCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTC	641
Db	551	HCFPGLGADCAKAAACPVLCSGNGQYSGTKQCQCYSGWKGAECDVPMNQCIDPSCGGHGSC	610
QY	642	ITGTICINPGYKGESCEEVDMDPTCSGRGVCVRGECHCFVGGTNCETPRATCLDQCS	701
Db	611	IDGNCVCSAGYKGEHCHEVDCLDPTCSSHGVCVNGECLSPGWGGLNCELARVQCPDQCS	670
QY	702	GHGTFLPDTGLCSCDPSWTGHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQRACH	761
Db	671	GHGTYLPDTGLCSCDPNMMGPDSCSVEVCSVDCGTHGVGICGACRCEEGWTGAACDQRVCH	730
QY	762	PRCAEHGTCRDGKCECSPGNWGEHCTIAHYLDRVVKEGCPGLCNGNRCRTLNLNGHVCVC	821
Db	731	PRCIEHGTCKDGKCECREGNGEHCTIGRTAGTETDGCPLDCLCNGNRCRTLGNQNSWCVC	790
QY	822	QLWREGAGCDTSMETACGSKDNDDGLVDCMDPDCCQLQPLCHINPLCLGSPNPLDIOE	881
Db	791	QTGWRGPGCNVAMETSCADNKNDEGLVDCLDPDCCCLQCSACGNSLLCRGSRDPLDIOQ	850
QY	882	TQVPVSQNLHSHFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNI	941
Db	851	GQ--TDWPAVKSFYDRIKLLAGKDSHTIIIPGENPENSLSVLIRGQVVTDTGTPLVGVNV	908
QY	942	SFVNNPLFGYTIISRODGSFGLVNTGGISIIILRERAPFITQEHTLWLPWDRFFVMEIIM	1001
Db	909	SPVKYPKYGYTITRQDGTDFDLIANGGASLTILHFERAPFMSQERTVWLPWNSFYAMDTLVM	968



gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;  
gene therapy; FCTR3a; neurestin-like protein; FCTR3f.

Homo sapiens.  
WO200166747-A2.  
13-SEP-2001.  
05-MAR-2001; 2001WO-US007160.  
03-MAR-2000; 2000US-0186592P.  
03-MAR-2000; 2000US-0186718P.  
06-MAR-2000; 2000US-0187293P.  
06-MAR-2000; 2000US-0187294P.  
17-MAR-2000; 2000US-0190400P.  
07-APR-2000; 2000US-0196018P.  
03-JAN-2001; 2001US-0259548P.  
(CURA-) CURAGEN CORP.  
Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;  
Macdougall J, Mishra V, Mezes PS, Rastelli L;  
XX  
WPI; 2001-596837/67.  
N-PSDB; AAS14089.  
Novel polypeptides designated as FCTR<sub>X</sub> polypeptides, useful in detection,  
prevention and treatment of a broad range of pathological states.  
Claim 1; Page 39; 215pp; English.  
The invention relates to human FCTR<sub>X</sub> polypeptides, FCTR1-FCTR7, and the  
nucleic acids encoding them. These sequences are useful for the treatment  
or prevention of numerous disorders including myelogenous leukaemia,  
carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal  
alloimmune thrombocytopaenia, neurological disorders, neurodegenerative  
disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie  
-Tooth neuropathy, demyelinating Gardner syndrome, familial  
myelodysplastic syndrome, mental health conditions, immunological  
disorders, allergy and infection, bronchial asthma, Avellino type  
eosinophilia, lung diseases, reproductive disorders, infertility, male  
and female reproductive disorders, deafness, glycoprotein Ia deficiency,  
desmoid disease, turcotte syndrome, liver cirrhosis, hepatitis C, gastric  
disorders, pancreatic diseases such as diabetes, Schistosoma mansoni  
infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,  
Corneal dystrophy-Greonow type I, Corneal dystrophy-lattice type I and  
Reis-Bucklers corneal dystrophy. This sequence represents FCTR3f, a  
homologue of FCTR3a protein

Sequence 2724 AA;

Query Match      64.0%;    Score 9530.5;    DB 4;    Length 2724;  
Best Local Similarity    61.4%;    Pred. No. 0;  
Matches 1755;    Conservative 414;    Mismatches 467;    Indels 221;    Gaps 26

QY    1    MDVKERKPYRLTRRR-DAERRYTSSSADSEEGKAP-QKSYSSETLKAYDQDARLAYGS 58  
     ||||| : : ||||| | : ||||| ||||| : : ||||| ||||| ||||| : : ||| :  
Db    1    MDVKDRR-HRSLTRGRCGKECRYTSSLDSDCRVPTQKSYSSETLKAYDHDSRMHYGN 59  
     ||||| : : ||||| | : ||||| ||||| : : ||||| ||||| ||||| : : ||| :  
QY    59   RVKDIVPQAEFEFCRTGANFTLRELGLEEVTPPHGTLYRTDGLPCQGYSMGAGSDADME 118  
     ||||| : : ||||| | : ||||| ||||| : : ||||| ||||| ||||| : : ||| :  
Db    60   RVTDLIHRSDEFFPRQGTNFTLAELGICEPS-PHRSGYCSDMGIHQYSLSTGSDADSD 118  
     ||||| : : ||||| | : ||||| ||||| : : ||||| ||||| ||||| : : ||| :  
QY    119   ADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHTETDHPG----- 168  
     : : ||||| : : ||||| | : ||||| ||||| : : ||||| ||||| ||||| : : ||| :  
Db    119   TEGGMSPEHAIRLWGRGIKRRSSGLSRENSALTLTDSDNENKSDDENGRPIPPTSPPS 178  
     : : ||||| : : ||||| | : ||||| ||||| : : ||||| ||||| ||||| : : ||| :  
QY    169   ----- 169  
Db    179   LLPSAQLPSSHNPVPVSCQMPLDNTSHQIMDTNPDEEFPSPSYLLRACSGPQQASSG 238  
QY    170   LQNH---ARLTTPPPPLSHAHTPNQHAAASINSLNRGNFTFRSN----PSPAPTDHLSLG 222

Db 1200 SLYVGDFNYIRIPPSRNVTSILELRNKEFKHSNNPAHKYLYLAVDPVSGSLYVSDTNSRR 1259

QY 1302 VFKIKSTVVVKDLVKNSEWAGTGDQCLPFDDTRCGDGGKATLTPNPRGITVDKFGLI 1361

Db 1260 IYRVKSLSGTDLAGNSEVAVAGTGEQCLPFDEARCGDGGKAIDATLMSPRGIAVDKNGLM 1319

QY 1362 YFVDGTMIRRIDONGIISTLLGSNDLTSARPLSCDSVMDISOVRLEWPTDLAINPMDNSL 1421

Db 1320 YFVDATMIRKVDONGIISTLLGSNDLTVAVRPLSCDSSMDVAQVRLEWPTDLAVNPMDSNL 1379

QY 1422 YVLDDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVATHATLESATALAVSHNGVL 1481

Db 1380 YVLENNVILRITENHQVSIAGRPMHCQVPGID-YLSKLAHSALESASAIASHTGVL 1438

QY 1482 YIAETDEKKINRQVTTSGEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSS 1541

Db 1439 YITETDEKKINLRQVTTNGEICLLAGASDCDCCKNDVNCNCYSGDDAYATDAILNSPSS 1498

QY 1542 LAVCADGELYVADLGNIRIRFIRKKNKPLNTQNMVELSSPIDBELYLFDDTTGKHLVTSQL 1601

Db 1499 LAVAPDGTIYIADLGNIRIRAVSKNKPVLNAFNQYEAASPGEBELVFNADGIHQYTVSL 1558

QY 1602 PTGDYLYNFTYTGDDITLITDNNGNMNVNRDSTGCMPLWLVPDGOVYVWVTMTGNSALK 1661

Db 1559 VTGEYLYNFTYSTDNDVTELIDNNGNSLKIIRDSSGMPRHLLMPDNQIITLTVGINGGLK 1618

QY 1662 SVTTQGHELAMTYHNGSGLLATKSNENGWTFYEYDSFGRLTNVTFPTGQVSSFRSDTD 1721

Db 1619 VVSTQNLELGLMTYDGTGLLATKSDBTGWTTFYDYDHEGRLTNVTRPTGVVTSLSHREME 1678

QY 1722 SSVHVQVETSSK-DDVTITTNLSASGAFYTLLOQDQVRNSYVIGADGSLRLLLLANGMEVAL 1780

Db 1679 KSITIDIENSRDDDDVTVITNLSSVEASYTVVQDQVRNSYQLCNGNGLTRVMYANGMGISF 1738

QY 1781 QTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVWRQORKEQARGQVTVFGRRLRVHNRNLLS 1840

Db 1739 HSEPHVLAGTITPTIGRCNISLPMENGLNSIEWRLRKEQIKGKVTIFGRKLRVHGRNLLS 1798

QY 1841 LDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTVSPGGYIAGIQGI 1900

Db 1799 IDYDRNIRTEKIYDDHRKFTLRILYDQVGRPFLWLPSSGLAAVNVSYFFNGRLAGLQGA 1858

QY 1901 MSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRYIFEFDKNDRLSSVTMPN 1960

Db 1859 MSERTDIDKQGRIVSRMFADGKWSYSYLDKSMVLLLSQSQRYIFEYDSSDRLLAVTMPSS 1918

QY 1961 VARQTLTIRSVGYVRNIYQPPENASVIQDFTEDGHLHLHTFYLGTRRRVIYKYGKLSKL 2020

Db 1919 VARHSMSTHTSIGYIRNIYNPPESNASVIFDYSDDGRILKTSFLGTGRQVYKYGKLSKL 1978

QY 2021 AETLYDTTKVSPTYDETAGMLKTNLQNEGFTCTIRYRQIGPLIDRQIFRFTHEGMVNAR 2080

Db 1979 SEIVYDSTAVTFGYDETTGVLKMNLSQSGFSTIRYRKIGPLVDKQIYRFSEEGMVNAR 2038

QY 2081 FDYNY-DNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITAVM 2139

Db 2039 FDYTHDNSFRIASIKPVISETPLPVDLYRYDEISGKVEHFGKFGVIYYDINQIITAVM 2098

QY 2140 THTKHFDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVKELKVGYPYANTTRYSEYDAD 2199

Db 2099 TLSKHFDTHGRIKEVQYEMFRSLMYMTVQYDSMGRVIRKELKLPYANTTKYIYDYDGD 2158

QY 2200 GOLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKXWDEDG 2259

Db 2159 GOLQSVAVNDRPTWRYSYDLNGLNHLNPGNSVRLMPLRYDLDRITRLGDVQYKIDDDG 2218

QY 2260 FLRQRGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGRVSSKSSHHLQFFYADLTN 2319

Db 2219 YLCQRGSDIFEYNSKGLLTRAYNKASGWSVQYRYDGVGRRASYSKTNLGHHLQYFYSDLHN 2278

QY 2320 PTKVTHLYNHSSEITSLYYDLQGHLPAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQ 2379

Db 2279 PTRITHVYNHSNSEITSLYYDLQGHLPAMESSSGEYYVASDNTGTPLAVFSINGLMIKQ 2338

QY 2380 ILYTAYGEIYMDTNPNFQIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHKLWKLH 2439

Db 2339 LQYTAYGEIYYDSNPDFQMVIGFHGGLYDPLTKLVHFTQRDYDVLAGRWTSPDYTMWKNV 2398

QY 2440 SSSNVMPENLYMFKNNNPISNSQDIKCFMTDVNSWLLTFGFGQLHNVIPGYPKPMDMAMEP 2499

Db 2399 GKEPA-PFNLYMFKSNPNLSSSELDLKNYVTDVKSWLVMFGFQLSNIIIPGFPRAKMYFVPP 2457

QY 2500 SYELIHTQMKTOEWDNSKSIILGVQCEVQKQLKAFVTLERFDQLYGSTITSCQQAPKTKK- 2558

Db 2458 PYELSESQAS----ENGQLITGVQQTTERHNQAFMALE-----GQVITKKLHASIREKA 2507

QY 2559 ---FASSGSVFGKGVKFKALDGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVD 2615

Db 2508 GHWFATTPITIGKIMFAIKRGRVTTGVSSIASEDSRKVASVLNNAYYLDKMHYSIEGKD 2567

QY 2616 THYFVKPGPSEGDLAAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCL 2675

Db 2568 THYFVKIGSADGDLVTLGTTIGRKVLESGVNVTVSQPTLLVNGRTRRFTNIEFYQYSTLLL 2627

QY 2676 NTRYG---TTLDEEKARVLELARQRAVRQAWAREQQRRLREGBELRAWTEGEKQQVLSTG 2732

Db 2628 SIRYGLTPDTLDEEKARVLDQARQALGTAWAKEQQKARDGREGSRLWTEGEKQQLLSTG 2687

QY 2733 RVQGYDGFVFVISVEQYPELSDSANNIHFMRQSEMGR 2769

Db 2688 RVQGYEGYVLPVEQYPELADSSNIQFLRQNMGMGR 2724

RESULT 11

ADB32029

ID ADB32029 standard; protein; 2724 AA.

XX ADB32029;

XX 04-DEC-2003 (first entry)

XX Human FCTR3f polypeptide.

DE Human; FCTR3f; colorectal cancer; adenomatous polyposis coli;

XX myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;

XX ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;

XX renal cell carcinoma; melanoma; clear cell carcinoma;

XX granular cell carcinoma; neurological disorder;

XX neurodegenerative disorder; nerve trauma;

XX familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;

XX Gardner syndrome; mental health condition; immunological disorder;

XX allergy; asthma; lung disease; reproductive disorder; deafness;

XX glycoprotein deficiency; desmoid tumour; turcot syndrome;

XX liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;

XX diabetes; schistosoma mansoni infection; spinocerebellar ataxia;

XX plasmodium falciparum infection; Groenouw's corneal dystrophy;

XX lattice corneal dystrophy.

XX Homo sapiens.

PN US2003087816-A1.

PD 08-MAY-2003.

XX 05-MAR-2001; 2001US-00800198.

XX 03-MAR-2000; 2000US-0186592P.

XX (VERM/) VERMET C.

PA (FERN/) FERNANDES E.

PA (SHIM/) SHIMKETS R.

PA (HERR/) HERRMANN J.

PA (MAJU/) MAJUMDER K.

PA (MACD/) MACDOUGALL J.

PA (MISH/) MISHRA V.

PA (MEZE/) MEZES P S.



Db 1439 YITETDEKINRLRQVTTNGEICLLAGAADCDCKNDVNCNCYSGDDAYATDAILNSPSS 1498  
QY 1542 LAVCADGELYVADLGNIRIRFIRKKNPFLNTQNMVELSSPIDOELYLFDTTGKHLYTQSL 1601  
Db 1499 LAVAPDGTIYIADLGNIRIRAVSKNKPVLNAFNQYEAASPGEOQLYVENADGIHQYTVSL 1558  
QY 1602 PTGDYLYNFTYTGDGITLITDNNGNMNVRRDSTGMPWLVLVDPDQVYVWVTGNTSALK 1661  
Db 1559 VTGEYLYNFTYSTDNDVTLEDIDNNGNSLKIRRDSSGMPRHLLMPDNQIITLTVTGNGGLK 1618  
QY 1662 SVTTQGHELAMMTYHNSGLLATKSNENGWTFYFYEDSFGRLLTNVTFPTGOVSSFRSDTD 1721  
Db 1619 VVSTQNLGLMTYDNGTGLLATKSDETGWTFTFYDYDHEGRLTNVTRPTGVVTSLHREME 1678  
QY 1722 SSVHVQVETSSK-DDVTITTNLSASGAFYTLLOQVNRNSYYIGADGSLRLLLLANGMEVAL 1780  
Db 1679 KSITIDIENSNRDDVTIVITNLSVEASTYVVQDQVRNSYQLCNNGTLRVMYANGMGISF 1738  
QY 1781 QTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEMRQKEQARGQVTVFGRRRLRVHNRNLS 1840  
Db 1739 HSEPHVLAGTITPTIGRCNISLPMENGLNSIEWRLRKEQIKGKVTIFGRKLRVHGRNLS 1798  
QY 1841 LDFDRVTRTEKIYDHRKFTLRILYDQAGPPLWSPPSSRLNGVNVVYSPGGYIAGIQRI 1900  
Db 1799 IDYDRNIRTEKIYDHRKFTLRILYDQVGRPFLWLPSSGLAANVSYFFNGRLAGLQGA 1858  
QY 1901 MSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLLSSVTMPN 1960  
Db 1859 MSERTDIDKQGRIVSRMFADGKWSYSLDKSMVLLQSQRQYIFEYDSSDRLLAVTMPS 1918  
QY 1961 VARQTLETIRSVGYRNIYQPPEGNASVIQDFTEDGHLHHTFYLGTRRRVIYKYGKLSKL 2020  
Db 1919 VARHSMSTHTSIGIRNIYNPPESNASVIFDYSDDGRILKTSFLGTGRQVYKYGKLSKL 1978  
QY 2021 AETLYDTTKVSFTYDETAGMLKLTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGMVNAR 2080  
Db 1979 SEIVYDSTAVTFGYDETGTGLKVNLSQSGFSCTIRYKIGPLVDKQIYRFSEEGMVNAR 2038  
QY 2081 FDYNY-DNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINOIITAVM 2139  
Db 2039 FDYTHDNSFRASIKPVISETPLPVDLYRYDEISGKVEHFGKFGVIYYDINOIITAVM 2098  
QY 2140 THTKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNMGVVKKELKVPYANTTYSYEYDAD 2199  
Db 2099 TLSKHFDTHGRIKEVQYEMFRSLMYWMTVQYDSMGRVIKRELKLGPYANTTKYTYDYDGD 2158  
QY 2200 GOLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDGD 2259  
Db 2159 GOLQSVAVNDRPTWRYSYDLNGLNHLNPGNSVRLMPLRYDLDRITRLGDVQYKIDDDG 2218  
QY 2260 FLRQRGGDIFEVNSAGLLIKAYNRAGSVRYRYDGLGRRVSKSSHHLQFFYADLTN 2319  
Db 2219 YLCQRGSDIFEVNSKGLLTRAYNKASGVSQYRYDGVGRRASYKTNLGHHLQYFYSDLHN 2278  
QY 2320 PTKVTHLYNHSSEITSLYYDLQHLFAMELSSGDEFYIACDNIGTPTLAVFSGTGLMIKQ 2379  
Db 2279 PTRITHVYNHNSSEITSLYYDLQHLFAMESSSGEEYVASDNTGTPLAVFSINGLMIKQ 2338  
QY 2380 ILYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHLMKHL 2439  
Db 2339 LQYTAYGEIYDNPDPFQMVIGFHGGLYDPLTKLVHFTQRDYDVLAGRWTSPDYTMWKNV 2398  
QY 2440 SSSNVMPFNLNFKNNPNISNSQDIKCFMTDVNSWLLTFGFOLHNVIPGYPKPDMDAMEP 2499  
Db 2399 GKEPA-PFNLNFKSNPNLSSELDLKNVYTDVKSWLVMFGQLSNIPGFPRAKMYFVPP 2457  
QY 2500 SYELIHTQMKTOBWDNSKSIILGVOCEVQKQKAFVTLERFDQLYGSTITSCQQAPKTKK- 2558  
Db 2458 PYELSESQAS----ENGQLITGVQQTTERHNQAFMALE-----GQVITKKLHASIREKA 2507  
QY 2559 ---FASSGSVFGKGVKFKALDGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVD 2615  
Db 2508 GHWFATTTPIIGKIMFAIKEGRVTTGTGVSSIASEDSRKVASVLNNAYYLDKMHYSIEGKD 2567

QY 2616 THYFVKPGPSEGLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCL 2675  
Db 2568 THYFVKIGSADGDLVTLGTTIGRKVLESGVNVTVSQPTLLVNGRTRRFTNIEFYSTLL 2627  
QY 2676 NTRYG---TTLDEEKARVLELARQAVRQAWAREQQRLREGEGEGLRAWTEGEKQQVLSTG 2732  
Db 2628 SIRYGLTPDTLDEEKARVLDQARQALGTAWAKEQQKARDREGSRLWTEGEKQQLLSTG 2687  
QY 2733 RVQGYDGGFFVISVEQYPELSDSANNIHFMRQSEMGR 2769  
Db 2688 RVQGYEGYVLPVEQYPELADSSSNIQFLRQNMGR 2724  
RESULT 12  
ABG61913  
ID ABG61913 standard; protein; 2725 AA.  
XX AC ABG61913;  
XX DT 15-AUG-2002 (first entry)  
XX Prostate cancer-associated protein #114.  
DE Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
XX OS Mammalia.  
XX PN WO200230268-A2.  
XX PD 18-APR-2002.  
XX PF 12-OCT-2001; 2001WO-US032045.  
XX PR 13-OCT-2000; 2000US-00687576.  
PR 08-DEC-2000; 2000US-00733288.  
PR 08-DEC-2000; 2000US-00733742.  
PR 24-JAN-2001; 2001US-0263957P.  
PR 16-MAR-2001; 2001US-0276791P.  
PR 16-MAR-2001; 2001US-0276888P.  
PR 06-APR-2001; 2001US-0281922P.  
PR 24-APR-2001; 2001US-0286214P.  
PR 30-APR-2001; 2001US-00847046.  
PR 04-MAY-2001; 2001US-0288589P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX PA Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX PI WPI; 2002-471335/50.  
XX DR N-PSDB; ABK92230.  
PT Detecting a prostate cancer-associated transcript in a cell in a patient,  
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
PT by determining if prostate cancer-associated genes are expressed in a  
PT prostate tissue.  
XX Claim 27; Page 397; 436pp; English.  
XX The present invention relates to methods of detecting a prostate cancer-  
XX associated transcript in a cell from a patient. The method comprises  
XX contacting a biological sample from the patient with prostate cancer-  
XX associated polynucleotides (designated PC genes) that selectively  
XX hybridise to a sequence that is at least 80% identical to them. The  
XX prostate cancer-associated polynucleotide sequences are differentially  
XX expressed in prostate tumour tissue or in prostate cancer and are derived  
XX from the tissues of various organisms such as humans or other mammals  
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for  
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-  
XX associated genes are useful for diagnosing or treating prostate cancer,  
XX as well as for identifying modulators of prostate cancer or agents that  
XX inhibit prostate cancer. The nucleic acid sequences are particularly  
XX useful in gene therapy, as a vaccine or in antisense applications.

CC	ABG61800-ABG61944 represent prostate cancer-associated proteins
XX	
SQ	Sequence 2725 AA;
	Query Match 63.8%; Score 9501; DB 5; Length 2725;
	Best Local Similarity 62.0%; Pred. No. 0;
	Matches 1737; Conservative 404; Mismatches 550; Indels 112; Gaps 17;
QY	1 MDVKERKPYRSLTR-RRDAERRYTSSSADSEEGKAPQKSYSSSETLKAVDODARLAYGS- 58
Dd	1 MEQTDCKPYQPLPKVKHEMDLAYTSSDSEDEGKRPQSYNSRETLHEYNQELRMNYSQ 60
QY	59 --RVKDIVPQEA-EFCRTGANFTLRELGLEEVTPHGTLYRTDI-GLPQCGYSMGAGSD 114
Dd	61 SRKRKEVEKSTQEMEFCEFSHTLC-----SGYQTDHMSVSRHGYLEMGSD 106
QY	115 ADMEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLITLDTHEHENTEDHPGGLQ--- 171
Dd	107 VDTETEGAASPDHALRMWIRGMKSEHSSCLSSRANSALSITDTHER-KSDGENGFKFSP 165
QY	172 -----NHARLR--TPPPPLSHAHTPNQHHAAASINSLNRGNFTPRS 209
Dd	166 VCCDMEAQAGSTQDVQSSPHNQFTFRPLPPPPPPHACTCARKPPPPAADSLQRRSMTTRS 225
QY	210 NPSAPPTDHSLSGEPAGGAQEPAAHAQENWLLNSNIPLLETRNLGKQPFLGTLQDNLIEMD 269
Dd	226 QPSPA-----APAPPTSTQDSVHLHNSWVLSNIPLETR----- 259
QY	270 ILGASRHDGAYSDGHFLFKPG-GTSPFLFCTTSPGYPLTSSTVSYSPPPRPLPRSTFARPAF 328
Dd	260 -----HSLFXHSGSGSSAIFSAASQNYPLTSTNTVSYSPPPRPLPRSTFSRPAF 305
QY	329 NLKKPSKYCNWKAALSALVISATLVILLAYFVAMHLFGLNWHLQPMEGQMYEITEDTAS 388
Dd	306 TENKPYRCNWKCTALSATAITVTLALLAYVIAVHLFGLTWQLQPVVEGELYANGSVKGN 365
QY	389 SWBVPTDVSLYPSGGTGLETPDRKKGTTTEGKPSFFPEDSFIDSGEIDVGRRASQKIPP 448
Dd	366 RGTESMDTTYSPIGKVSXDKSEK-----VFQKGRAIDTGEVDIGAQMQTIPP 414
QY	449 GTFWRSQVFIIDHPVHLKFNYSGLGAALVGIYGRKGLPPPSHTQDFVELLDGRLLTQEAR 508
Dd	415 GLFWRFQITIHPIYLVKFNISLAKDSLGIYGRRNIPPTHQTFDFVKLMDGKQLVKQDSK 474
QY	509 SLEGTPRQSRGTVPSPSSHETCFIQYLDSGIWHLAFVNDGKSESVSFLTTATIESVDNCPS 568
Dd	475 GSDDTQHSRPNLILTSLOETGFIYMDQGPWYLAFYNDGKKMEQVFVLTITTAIEIMDDCST 534
QY	569 NCYNGGDCISGTCHCFGLGFPDCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECDVPTN 628
Dd	535 NCNGNGECISGHCHCFPGFLGPDPCARDSCPVLCCGNGEYEGKHCVCVRHGWKGPECDVPEE 594
QY	629 QCIDVACSNHGTICITGTCINPGYKGESCEEVDGMBTCSGRGVCVRGECHCFVGGWGTN 688
Dd	595 QCIDPTCFGHGTCIMGVCI CVPGYKGEICEEEDCLDPMCSNHGICVKGECHCSTGWGGVN 654
QY	689 CETPRATCLDQCSEGHGTFLPDTGLCSCDPSWTGHDCSI EICAADCGGHGVCVCGTCRCED 748
Dd	655 CETPLPVCQEQCSGHGTFLLDAGVCSDDPKWTGSDCSTELCTMECGSHGVCISRGICQCEE 714
QY	749 GNMGAACDQRACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNG 808
Dd	715 GWVGTEERSCHSHCTEHGQCKDGKCECSPGWEGDHCTIAHYLD-AVRDGCPLCFGNG 773
QY	809 RCTLDLNGWHCVQCQLGWRGAGGDTSMETACGDSKDNDDGLVDCMDPDCCLOPLCHINPL 868
Dd	774 RCTLDQNGWHCVQCQVGSWTGCVNVVMEMLCGDNLNDNDGDLTDCVDPCCQCSNICYISPL 833
QY	869 CLGSPNPLDIIQETQVPVSQQNLHFSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOV 928
Dd	834 CQGSPPDLDIQQSQTLFSQHTSRLFYDRIKFLIGKDSHTVIPPEVSFDSRRACVIRGOV 893
QY	929 MTSDGTPLVGNISFVNNPLFCYTISRQDGSFDLVTNGGISILRFRERAPFITQHTLWL 988

Dd	894 VADGTPLVGNVSFLHHSYDGTISRQDGSFDLVAIGGISVILIFDRSPFLPEKRTLWL 953
QY	989 PWDREFFVMEITIIIMRHEENEIPSCDLSNFARPNPVVSPPLTSPASSCAEKGPIVPEIQAL 1048
Dd	954 PWNQPIVVEKVTMQRVVSDPSCDISNFISPNPIVLPSPLTSPGSGCSPERGTTIVPELQVV 1013
QY	1049 QEEISISGCKMRLSYLSSRTPGYKSVLRISLTHPTTIPNLMKVHLMVAVEGRLEFRKWFPA 1108
Dd	1014 QEEIPIPSSFVRLSYLSSRTPGYKTLRLILLTHSTIPVGMIKVHLTVAVEGRLLTKQWFFA 1073
QY	1109 APDISYFIWDKTDVYNQVFGLSAEAFVSUGYEYESCDLILWEKRTTIVLQGYEIDASKL 1168
Dd	1074 AINLYTFAWNKTDIYGQKVWGLAEALVSGYEYETCPDFILWEQRTTVLQGFEMDASNL 1133
QY	1169 GGWSLDKHHALNIQSGILHKNGENQFVSQQPPVIGSIMNGRRRSISCPSCNGLADGNK 1228
Dd	1134 GDWSLNKHHILNPQSGIIHKNGENMFISQPPVISTIMGNHGQRSVACTNCNGPAHNK 1193
QY	1229 LLAPVALTCGSDGSLYVGFNVIIRRIFFPSGNVTNILELNKDFRHSHPAHKYIYLATDPM 1288
Dd	1194 LFAPVALASGPDGSVYVGFNVRIRFPFSGNSVSILEL-----STSPAHKYIYLAMDPV 1246
QY	1289 SGAVFLSDNSRRVFKIKSTVVVKDLVKNSEVAGTGDQCLPFDDTRCGDGGKATEATLT 1348
Dd	1247 SESLYSDTNTKVKYKLSLVETKDLKSNFEVVAGTGDQCLPFQSHCGDGRASEASLN 1306
QY	1349 NPRGITVDKFLIYFVDGTMIRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEW 1408
Dd	1307 SPRGITVDRHGFIYFVDGTMIRKIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVRLEW 1366
QY	1409 PTDLAINPMDNSLYVLNINVLIQISENHQVRIVAGRPVHCQVPGIDHFLLSKVAIHATLE 1468
Dd	1367 PTDLAVNPMDNSLYVLNINVLIQISENRRVRIIAGRPIHCQVPGIDHFLVSKVAIHSTLE 1426
QY	1469 SATALAVSHNGVLYIAETDEKKINRIQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDD 1528
Dd	1427 SARATSVSHSGLLFIAETDERKVNRIQVTTNGEYIYIACAPTDCKIDPNCDCFSGDG 1486
QY	1529 GYAKDAKLNTPSSLAVACDAGELYVADLGNIRIRFIRKKNPFLNTQNMVELSSPIDQELYL 1588
Dd	1487 GYAKDAKMKAPSSLAVSPDGILYVADLGNVRIRTIISRNOAHLNMDNMIYEIASPADQELYQ 1546
QY	1589 FDTTGKHLTYQSPLPTGDYLYNFYTGDDGITLITDNNGNMNVNRRDSTGMPLWLVPDQG 1648
Dd	1547 FTVNGTHLHLNLITRDYVYNYFTYNSGDLGAIITSSNGNSVHIRRDAGGMPLWLVPDGG 1606
QY	1649 VYVVTMGTNSALKSVTTQGHELAMTYHGNISGLLATKSNENGTTFEYEDSFGRLTNVTF 1708
Dd	1607 VYWLTISSNGVLKRVSQAQYNPALMTYPGNTGLLATKSNENGTTVVEYDPEGHLTNATF 1666
QY	1709 PTQVSSFRSDTSSVHVQVETSSKDDVTITTNLSASGAFYTLQDQVRNSYYIGADGSL 1768
Dd	1667 PTGEVSSPHSDLEKLTKVELDTSNRENVLMSNTLTATSTIYILKQENTQSTYRVNPDGSL 1726
QY	1769 RLLLANGMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEWRQRKEQARGQVTVFG 1828
Dd	1727 RVTFASGMEIGLSSEPHILAGAVNPTLGKCNISLPGEHNANLIEWRQRKEQNKGNVSAFE 1786
QY	1829 RRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSSLWSPSSRLNGVNVITYS 1888
Dd	1787 RRLRAHNRNLLSIDFDHITRTGKIYDDHRKFTLRILYDQTRPILWSPVSRVNEVNITYS 1846
QY	1889 PGGYIAGIQRGIMSERMEYDQAGSITSRIFADGKTWSYTYLEKSWVLLHLSORQYIFEDF 1948
Dd	1847 PSGLVTFIQRGITWNEKMEYDQSGKIISRTWADGKIWSYTYLEKSWVLLHLSORRYIFEYD 1906
QY	1949 KNDRLLSVTMPNVARQTLTIRSVGYRYNIYQPEGNASVIQDFTEDGHLHHTFYLGTGR 2008
Dd	1907 QSDCLLSVTMPMSMVRHSLQTMLSVGYYRYNIYTPPDSSTSFQDYSRDRGLLQTLHLGTGR 1966
QY	2009 RVIIKYGKLSKLAETLYDTTKVSFTYDETAGMLKTLINLQNEGFTCTIRYRQIGPLIDRQI 2068

Db 1967 RVLYKYTKQARLSEVLDTTQVTLTYEESGVIKTIHLMHDGFICTIRYQGTGLIGRQI 2026  
QY FRPTEEGMVARFVYDNSPRVTSMAQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYY 2128  
Db 2027 FRPSEGLVNAFVYSY-NNFRVTSMAQAVINETPLPIDLYRYDDVSGKTEQFGKFSVINY 2085  
QY DINQIITAVMTHYKHFDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVKKELKVGYPYAN 2188  
Db 2086 DLNQVITTTVMKHTKIFSANGQVIEVQYELKAIAYWMTIQYDNVGRHGNMCIRVGVDAN 2145  
QY TTRYSEYDADGQLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRL 2248  
Db 2146 ITRYFYEYDADGQLQTVSVNDKTQWRYSYDLNGLDINLLSHGKSARLTPLRYDLRDRITRL 2205  
QY GDVQYKMDGFLRQGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGRRVSSKSSH 2308  
Db 2206 GEIQYKMDGFLRQGGDIFEYNSAGLLIKAYNRAGSWTVQYDYDGLGRRVASKSLGQ 2265  
QY HLQFFYADLTNPVKVTHLYNHSSEITSLYDYLQGHFLFAMELSSGDEFYIACDNIGTPLA 2368  
Db 2266 HLQFFVDATEANPIRVTHLYNHTSSEITSLYDYLQGHFLFAMELSSGEEYVACDNTGTPLA 2325  
QY VFSGTGLMIKQILYAYGEIYMDTNPFNQIIGYHGLYDPLTKLVHMGRRDYDVLGSRW 2428  
Db 2326 VFSRGQVKEILYTPYGDYHDTYDFQVIGFHGGLYDFLTKLVHLGQRDYDVVAGRW 2385  
QY TSPDHELKHLSSNMV--PFLNLYMFKNNPISNSQDIKCFMTDVNSWLLTFGFQHLNVI 2486  
Db 2386 TTAYHIIWKQL---NLLPKFNLVSFENYVPGKIQDVAKYTTDIRSWLELFGFQHLNVL 2442  
QY PGYPKPDMDAMEPSYELIHTQMKTOEWDNSKSLGVQCEVQKQKAFVTLERFDQLYGST 2546  
Db 2443 PGFPKPELENLELYELLRLQTKTQEWDPKTLGIQCELOKQLRNFIQLDQLPMTPRYN 2502  
QY ITSCQAPKTKKFPASSGSVFGKGVKFAKDKGRVTTDIISVANEDGRRVAAAILNHAHYLEN 2606  
Db 2503 DGRLEGKQKPRFAAVPSVFGKGIKFAKDKGIVTADIIGVANEDSRRLAAILNHAHYLEN 2562  
QY LHFTIDGVDTHYVFKPGPSEGLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDI 2666  
Db 2563 LHFTIEGRDTHYFIKLSLEEDLVLIQNTGRRILENGVNVTVSQMTSLNGRTRRFADI 2622  
QY QLQYGALCLNTRYGTTLDEEKARVLELARQAVRAQVAREQORLEGEGLRAWTEGEKQ 2726  
Db 2623 QLQHGALCFNTRYGTTVEEKNHVLARQAVRAQVATKEORLQEGEGEIRAWTEGEKQ 2682  
QY QVLSTGRVQGYDGFVVISVEQYPELSDSANNIHFMRQSEMGR 2769  
Db 2683 QLLSTGRVQGYDGFVLSVEQYLELSDSANNIHFMRQSEIGRR 2725  
RESULT 13  
ABR58318  
XX ABR58318 standard; protein; 2633 AA.  
AC ABR58318;  
XX  
DT 08-JUL-2003 (first entry)  
XX  
DE BC00205B protein #SEQ ID 82.  
XX  
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;  
KW drug discovery; clinical medicine; forensic medicine; chromosome 5q33.3.  
XX Homo sapiens.  
OS  
XX WO2003029421-A2.  
PN  
XX  
PD 10-APR-2003.  
XX  
PF 02-OCT-2002; 2002WO-US031287.  
XX  
PR 03-OCT-2001; 2001US-0326526P.

PR 14-MAY-2002; 2002US-00144194.  
XX (ORIG-) ORIGENE TECHNOLOGIES INC.  
XX  
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;  
XX  
DR WPI; 2003-381623/36.  
DR N-PSDB; ACC72052.  
XX  
PT New isolated human differentially-regulated breast cancer polynucleotide  
PT and polypeptide, useful for diagnosing, staging, prognosticating,  
PT preventing and/or treating diseases and conditions relating to breast  
PT cancer.  
XX  
PS Claim 6; SEQ ID NO 82; 127pp + Sequence Listing; English.  
XX  
CC The invention relates to isolated polynucleotides which are  
CC differentially-regulated in breast cancer. The methods and compositions  
CC of the present invention are useful for detecting, diagnosing, staging,  
CC monitoring, prognosticating, preventing and/or treating diseases and  
CC conditions relating to breast cancer, and may be used in gene therapy or  
CC antisense therapy. They can also be used in research, drug discovery,  
CC clinical medicine and forensic medicine. Sequences given in records  
CC ABR58278-ABR58323 and ABR58346-ABR58362 represent polypeptides encoded by  
CC polynucleotides of the invention that are differentially-regulated in  
CC breast cancer. NOTE: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2633 AA;  
Query Match 62.6%; Score 9314; DB 6; Length 2633;  
Best Local Similarity 61.9%; Pred. No. 0;  
Matches 1724; Conservative 414; Mismatches 477; Indels 172; Gaps 26;  
QY 1 MDVKERKPYRSLTRRR-DAERRYTSASDSEEGKAP-QKSYSSSETLKAYDQDARLAYGS 58  
Db 1 MDVKDRR-HRSLTRGRCGKCRYTSSSLDSEDCRVPTQKSYSSSETLKAYDHSRMHYGN 59  
QY 59 RVKDIVPQEAHEFCRTGANFTLRELGLEEVTPPHGTLRYRTDGLPQCGYSMGAGSDADME 118  
Db 60 RVTDLIHRSEDEFPRQGTNFTLAEIGICEPS-PHRSYGCSDMGLHQGYSLSGTSDADSD 118  
QY 119 ADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHAR--L 176  
Db 119 TEGGMSPEHAIRLWGRGIKRRSSGLSSRENSALTLTDSNENKSDDENGGPPNHHQSSTL 178  
QY 177 RTPPPPLSHAHTPNQOHAASINSLNENGNFTPRSN---PSAPTDSLSEPPAGGAQEP 232  
Db 179 RPPLPP-PHNHTLSHHH-SSANSLNRNSLTNRRSQIHAPAPAPND--LATTP-----ES 228  
QY 233 AHAQENWLLNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHDGAYSDFHFLFX-PGG 291  
Db 229 VQLQDSWVLSNVPLETR-----HFLFKTSSG 255  
QY 292 TSPLFCTTSPGYPLTSTVYSPPPRPLPRSTFARPAFNLKPKSKYCNWKAALSATVISA 351  
Db 256 STPLFSSSPGYPLTSGTVYTPPPRLLPRNTFSRKAFKLKPKSKYCSWKAALSATVISA 315  
QY 352 TLVILLAYFVAMHLFGLNWHLQNMEGQMEITEDETASSWPVPTDVSLYPSGGTGLETPDR 411  
Db 316 LLAILLAYFIAMHLGLNWLQPADGHTF--NNGIRTGLPGNDVATMPSSGK----- 366  
QY 412 KGKGTTEGKPSFFPDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHPVHLKFNVSIG 471  
Db 367 -----VPWSL--KNSSIDSGEAEVGRVTVQEVPPGVFWRSQVHIHSQPQFLKFNISLG 416  
QY 472 KAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEARSLBEGTPRQSRGTVPSSSHETGFI 531  
Db 417 KDALFGVYIRRGGLPPSHAQYDFMERLDGK-----EKWSVVEPRRRSIQTLVQNEAVFV 471  
QY 532 QYLDSGIWHLAFYNDGKESEVVSFLTITAESVDNCPNSYNGDCTISGTCFLGLGPD 591

Db 472 QYLDVGLWHLAFYNDGDKEMVSFNTVVLDVQDCPRNCHNGNCEVSGVCHCFPGFLGAD 531

QY 592 CGRASCPVLCSGNGQYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTITGTCICNPG 651

Db 532 CAKAACPVLCSGNGQYSGTKQCYSYSGWKGAECDDVPMNQCIDPSCGGHSGCIDGNCVCSAG 591

QY 652 YKGESCEEVDCMDPTCSGRGVGVGRGECHCFVWGVTNCETPRATCLDQCSCGHGTFPLPDIG 711

Db 592 YKGEHCEEVDCLDPTCSSHGVCVNGECLCPGWGGLNCELARVQCPDQCSGHGTYLPLDITG 651

QY 712 LCSCDPSWTGDCSTIEICAADCGGHGVGVGTGTCRCEGDMWGAACDQACHPRCAEHGTCR 771

Db 652 LCSCDPNMMGPDCSV----- 666

QY 772 DGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHVCVQQLGWRGAGCD 831

Db 667 -----DGCPCDLCNGNGRCTLGQNSWQCVQCTGWRGPGCN 700

QY 832 TSMETACGDSKNDGDLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPSVQQNL 891

Db 701 VAMETSCADNKDNEGDLVDCLDPDCCLOSAQCNLSLLCRGSRDPLDIIQQGQ--TDWPAV 758

QY 892 HSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOVMTSDGTPLVGVNISFVNNPLFGY 951

Db 759 KSFYDRIKLLAGKDSTHIIIPGENPFNSSVLIRGOVTTDGTPLVGVNVSVFKPKYGY 818

QY 952 TISRQDGSFDLVTNGGISIILRPERAPFITQEHTLWLPWDRFFVMTIIMRHEENEIPSC 1011

Db 819 TITRQDGTFDLIANGGASLTLHFERAPFMSQBERTVWLPWNSFYAMDILVMKTEENSIPSC 878

QY 1012 DLSNFARPNVPSPLTSPASSCAEKGPVIBEIQALQEEISISGCKMRLSYLSSRTPGY 1071

Db 879 DLSGFRVPDPPIIISSPLSTFFSAAPGQNPVIVPETQVHLHEEIELPGSNVVKLRYLSSRTAGY 938

QY 1072 KSVLRISLTHPTIPFNLMKVHLMVAVEGRLFRKWFAPADLSYFYFWDKTDVYNQKVFLG 1131

Db 939 KSLKKTMTQSTQVPLNLRVHLMVAVEGHLFQKSFQASPNLAYTFIWDKTDAYGQRVYGL 998

QY 1132 SEAFVSVGVEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKXHALNIQSGILHKNG 1191

Db 999 SDAVSVGFYEYETCPSLILWEKRTALLQGFELDPNGLGWSLDKXHLNVKSGILHKGTG 1058

QY 1192 ENQFVSQPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYI 1251

Db 1059 ENQFLTQOPAIITSIMNGRRRSISCPSCNGLAEGNKLAPVALAVIGIDGSLYVGDFNYI 1118

QY 1252 RRIFFSGNVTNILELRNKDPRHSHSPAHKYYLATDPMGSAVFLSDSNSRRVFKIKSTVVV 1311

Db 1119 RRIFFSRNVTSIELRNKEFKHSNPNPAHKYYLAVDPVSGSLYVSDTNSRRIYRVKSLSGT 1178

QY 1312 KDLVKNSEVVAGTGDCQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIR 1371

Db 1179 KDLAGNSEVVAGTGEQCLPFDEARCGDGKAIATILMSPRGIAVDKNGLMYFVDATMIRK 1238

QY 1372 IDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVVLQ 1431

Db 1239 VDQNGIISTLLGSNDLTAVRPLSCDSSMDVAQVRLEWPTDLAVNPMDNSLYVLENNVILR 1298

QY 1432 ISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLVYIAETDEKKI 1491

Db 1299 ITENHQVSIIAGRPMHCQVPGID-YLSKLAIHSALESASAIASHTGVLYITETDEKKI 1357

QY 1492 NRIRQVTSGEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLNTPSSLAYCAGDGELY 1551

Db 1358 NRLRQVTNGEICLLAGAASDCCKNDVNCNCSYSGDDAYATDAILNSPSSLA VAPDGTIY 1417

QY 1552 VADLGNIRIRFIRKNKPFNLNTQNMWYELSSPIDQELYLEFDTTGKHLTYQSLPTGDYLYNFT 1611

Db 1418 IADLGNIRIRAVSKNKPVLNAFNQYEAASPGQEQLYVFNADGIHQYTVSVLVTGELYLYNFT 1477

QY 1612 YTGCDGITLITDNNGMNVNRDSTGMPLWLVPDGVGVVVTMTGNTSALKSVTTQGHELA 1671

Db 1478 YSTDNDVTELIDNNGNSLKIRDSGMPRHLMPDNQIITLTVTGNGGLKVVSVSTQNLELG 1537

QY 1672 MMTYHGNSGLLATKSNENGWTTFFYEYDSFGRLTNVTFTPTGOVSSFRSDTSSVHVQVETS 1731

Db 1538 LMTYDNGTGLLATKSDBTGTTFFDYDHEGRLTNVTPTGVVTSLHREMEKSITIDIENS 1597

QY 1732 SK-DDVTITTNLSASGAFYTLQDQVRNSYIICADGSLRLLLLANGMEVALQTEPHLLAGT 1790

Db 1598 NRDDDVTVITNLSSVEASYTVVQDQVRNSYQLCNCNGTLRVMYANGMGISFHSBPHVLAGT 1657

QY 1791 VNPTVGKRNVTLPIDNGLNLVEWRQQRKEQARGQVTVFGRRLRVHNRNLLSLDPDRVTRTE 1850

Db 1658 IPTTIGRCNISLPMENGLNSIEWRLRKEQIKGKVTIFGRKLRVHGRNLLSIDYDRNIRTE 1717

QY 1851 KIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTPSPGGYIAGIQIGIMSERMEYDQA 1910

Db 1718 KIYDDHRKFTLRIIYDQGRPFLWLPSSGLAAVNVSYFFENGLRGLAGLQORGAMSERTDIDKQ 1777

QY 1911 GRITSRIFADGKTWSYTYLEKSMVLLLSHRSORQYIFEFDKNDRLSSVTMPNVVARQTLETIR 1970

Db 1778 GRIVSRMFADGKWSYSYLDKSMVLLLSQSORQYIFEYDSSDRLLAVTMPSPVARHSMSTHT 1837

QY 1971 SVGYRNIYQPPPEGNASVIQDFTEDGHLHLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKV 2030

Db 1838 SIGYIRNIYNPPESNASVIFDYSDDDRILKTSFLGTGRQVFKYKGLSKLSEIYVDSTAV 1897

QY 2031 SFTYDETAGMLKTINLQNEGFTCTIYRQIGPLIDRQIFRFTEEGMVNAREFVNY-DNSF 2089

Db 1898 TFGYDETTGVLKMNLSQSGFSCTIYRKIGPLVDKQIYRFSEEGMVNAREFYTYHDNSF 1957

QY 2090 RVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGYIYYDINQIITTAVMTHTKHFDAYG 2149

Db 1958 RIASIKPVISETPLPVLDLYRYDEISGKVEHFHFGKFGYIYYDINQIITTAVMTLSKHFDTHG 2017

QY 2150 RMKEVQYEIFRSLMYMTVQYDNMGRVVKELKVGYPVANTTRYSEYDADGQLQTVSIND 2209

Db 2018 RIKEVQYEMFRSLMYMTVQYDSMGRVIKRELKGLPYANTTKYTYDYGDLQQLQSVAVND 2077

QY 2210 KPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDGDFLRQRGDIF 2269

Db 2078 RPTWRYSYDLNGLNHLNPGNSVRLMPLRYDLRDRITRLGDVQYKIDDDGYLCQRGSDIF 2137

QY 2270 EYNSAGLLIKAYNRAGSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPNTKVTHLYNH 2329

Db 2138 EYNSKGLLTRAYNKASGWSVQYRYDVGRRASYKTNLGHHLQYFYSDLHNPTRITHVYNH 2197

QY 2330 SSSEITSLYYDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYATAGEIY 2389

Db 2198 SNSEITSLYYDLQHLFAMESSGEEYVVASDNTGTPLAVFSINGLMIKQLQYATAGEIY 2257

QY 2390 MDTNENFOIIGYHGGLYDPLTKLVHMGRRDYDVLAGRMTSPDHELKHLSSSNVMPFNL 2449

Db 2258 YDSNPDQMVIGFHGGLYDPLTKLVHFTQRDYDVLAGRMTSPDYTMWKNVKGKEPA-PFNL 2316

QY 2450 YMFKNPNPISNSQDIKCFMTDVNSWLLTFGQLHNVI PGYPKPDMDAMEPSYELIHTQMK 2509

Db 2317 YMFKSNNPLSSELDLKNYVTDVKSWLVMFGFQLSNIIPGEPRAKMYFVPPPYELSESQAS 2376

QY 2510 TQEWDNKSILGVQCEVQKQLKAPVTLERFDQLYSTITSCQQA PKTK-----FASSGSV 2565

Db 2377 -----ENGQITGVQQTTERHNQAFMALE-----GOVITKHLHASIREKAGHWFATTTPI 2426

QY 2566 FGKGVKFPALKDGRVTTDII SVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPS 2625

Db 2427 IGKIMFAIKEGRVTTGVSSIASEDSRKVASVLNNAYYLDKMHYSIEGKDTHYFVKIGSA 2486

QY 2626 EGDLAJLGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYG---TT 2682

Db 2487 DGDVLTLGTGTIGRKVLESQVNVTVSQPTLLVNGTRTRRTNIEFYQYSTLLSIRYGLTPDT 2546

QY 2683 LDEEKARVLELARQAVRQAWAREQORLREGEGELRAWTEGKQOQVLSTGRVQGYDGFFV 2742

Db 2547 LDEEKARVLDQARQALGTAWAKEQOKARDGREGSRLWTEGKQOQLLSTGRVQGYEGYV 2606



Qy	1404	VRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRIIVAGRPMHCQVPGIDHFFLLSKVAI	1466
Db	1189	VRLEWPTDLAVNPMDNSLYVLENNVILRITENHQVSIIVAGRPMHCQVPGID-YSLSKLAI	1247
Qy	1464	HATLESATALAVSHNGVLYIAETDEKKNIRIQVTTSGEISLVAGAPSGCDCKNDANCDC	1523
Db	1248	HSALESASAIASHTGVLYITETDEKKNIRLQVTTNGEICLLAGAASDCDCKNDWCNC	1307
Qy	1524	FSGDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRIRFIRKNKPFLLNTQNMVELSSPID	1583
Db	1308	YSGDDAYATDAILNSPSSLAVAPDGTIYIADLGNIRIRAVSKNKPVLNAFNQYEAASPE	1367
Qy	1584	QELYLFDTTGKHLTYQSLPTGDYLYNFYTGDDITLITDNNGNMNVNRDSTGMPLMLV	1643
Db	1368	QELYVFNADGHIQYTVSVLTGEYLYNFYTYSTDNDVTELIDNNGNSLKIRRDSSGMPRHLL	1427
Qy	1644	VPDGVYVWVTMGNSALKSVTTQGHELAMTYHNGSGLLATKSNENGWTTTFEYDSFGRL	1703
Db	1428	MPDNQIITLTVGNGGLKVSVSTQNLELGLMTYDGNLTGGLATKSDETWTTFYDYDHEGRL	1487
Qy	1704	TNVTFTQGVSSPESDTSVVHVQVETSSK-DDVTITTNLSASGAFYTLQDOVRNSYYI	1762
Db	1488	TNVT RTPGVVTSLHREMEKSIITIDIENSNRDDVTIVITNLSVEASYTVVQDOVRNSYQL	1547
Qy	1763	GADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEMRQKEQARG	1822
Db	1548	CNNGTLRVMYANGMISFHSHPVLAGTITPTIGRCNISLPMENGLNSIEWLRKEQIKG	1607
Qy	1823	QVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDHRKFTLRILLYDQAGRPSLWSPSSRLNG	1882
Db	1608	KVTIFGRKLRVHGRNLLSIDYDRNIRTEKIYDHRKFTLRIRIYDQGRPFPLWLPSSGLAA	1667
Qy	1883	VNVYSPGGYIAGIQRGIMSERMEYDQAGRITSIFADGKTWSYTYLEKSMVLLHLSQRQ	1942
Db	1668	VNVSYFFNGRLAGLQRGAMSERDIDKQGRIVSRMEFADGKVWSYLYLKDMSVLLQSQORQ	1727
Qy	1943	YIFEDKNDRLSSVTMPNVARQTLTIRSVGYVRNIYQPPEGNASVIQDFTEDCHLLHTF	2002
Db	1728	YIFEYDSSDRLLAVTMPSVARHSMSTHTSIGYIRNIYNPPESNASVIFDYSDDGRILKTS	1787
Qy	2003	YLGTRRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGP	2062
Db	1788	FLGTGRQVFKYIGKLSKLSIYVDSTAVTFGYDETTGVLMVNLQSGGFSCTIRYRKIGP	1847
Qy	2063	LIDRQIFRTEEGMVNARFDYNY-DNSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFG	2121
Db	1848	LVDKQIYRFSEEGMVNARFDYTYHDNSFRIASIKPVISETPLPVDLYRYDEISGVKREL	1907
Qy	2122	KFGVIYYDINQIITAVMTHTKHFDAYGRMKEVOYEIFRSLMYWMTVQYDNMGRVVKEL	2181
Db	1908	KFGVIYYDINQIITAVMTLSKHFDTHGRIKEVOYEMFRSLMYWMTVQYDSMGRVIKREL	1967
Qy	2182	KVGPYANTTRYSEYDADGQLQTVSINDKPLRWYSYDLNGNLHLLSPGNSARLTPLRYDI	2241
Db	1968	KLGPYANTTKYTYDGDGQLQSVAVNDRPTWRYSYDLNGNLHLLNPGNSVRLMPLRYDL	2027
Qy	2242	RDRITRLGDVQYKXMDDEGFLRQGGDIFEYNSAGLLIKAYNRAGSWSVRYYRVDGLGRRVS	2301
Db	2028	RDRITRLGDVQYKIDDDGYLCQGSDFEYNSKGLLTRYAYNKASGWSVQYRYDGVGRRAS	2087
Qy	2302	SKSSHSHHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQGHLFAMELSSGDEFYIACD	2361
Db	2088	YKTNLGHHLQYFYSDLHNPTRITHVYNHSNSEITSLYYDLQGHLFAMESSSGEEIYVASD	2147
Qy	2362	NIGTPLAVFSGTGLMIKQILYATAGEIYMDTNPNFQIIGYHGGLYDPLTKLVNMGRRDY	2421
Db	2148	NTGTPLAVFSINGLMIKQLOYATAGEIYYDSNPDEFQVMVIGFHGGLYDPLTKLVHFTQRDY	2207
Qy	2422	DVLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNPISNSQDIKCFMTDVNSWLLTFGFQ	2481
Db	2208	DVLAGRWTSPDYTMWKNVCKEPA-PFNLYMFKSNPNLSSSELDLKNYVTDVKSWLYMFGFQ	2266
Qy	2482	LHNVIPGYPKPDMAVEPSYELIHTQMKTOEWDNSKILGVQCEVQKQLKAFVTLIERFDQ	2541

Db	2267	LSNIIPGPRAKMYFVPPPYELSESQAS	---ENGQLITGVQQTTERHQAFALE---	2318
Qy	2542	LYGSTITSCQQAPKTKK	---FASGSVFGKGVKFAKDGRTTDDIISVANEDGRRVAAI	2597
Db	2319	--GQVITKKLHASIREKAGHWFAITPIIGKIMFAIKEGRVTTGVSSIASEDSRKVASV		2376
Qy	2598	LNHAHYLENLHFTIDGVDTHYFVKPCPSEGDLAILGLSGRRRTLENGVNVTVSQINTVLN		2657
Db	2377	LNNAVYLDKWHYSIEGKDTHYFVKIGSADGDLVTLTGTIGRKVLESGVNVTVSQPTLLVN		2436
Qy	2658	GRTRRYTDIQLQYGALCLNTRYG	---TTLDEEKARVLELARQRAVRQAWAREQQORLREGE	2714
Db	2437	GRTRRFTNIEFQYSTLLLSIRYGLTPDTLDEEKARVLDQARQALGTAWAKEQQOKARDGR		2496
Qy	2715	EGLRAWTEGEKQQVLSTGRVQGYDGFVVISVEQYPELSDSANNIHFMRSQSEMGRR		2769
Db	2497	EGSRLWTEGEKQQOLLSTGRVQGYEGYVYLPVEQYPELADSSNIQFURQNMCKR		2551

RESULT 15

AAM78695

ID AAM78695 standard; protein; 2136 AA.

AC AAM78695;

06-NOV-2001 (first entry)

XX	Human protein	SEQ ID NO 13
DE		

Human: cytokine: cell prolifera-

KW vaccine growth factor: immunomodulatory peptide therapy; stem

KW nervous system disorder; arthritis  
 YX

US  
yy  
Homo sapiens.

PN WO20015/190-A2.

PD 09-AUG-2001.

05-FEB-2001; 20

PR 03-FEB-2000;

PR 20-JUN-2000; 2000US-00598075.

PR 01-SEP-2000; 2000US-00654936.

PR 20-OCT-2000; 2000US-00693325.



PI Ma Y, Zhao QA, Wang D, Wan






































DR N-PSDB; AAK51828.

PT Nucleic acids encoding polype

3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527

2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526

CC encoded polypeptides (AAM78323-AAM80302) t

cytokine, cell proliferation of cells and production of other cytokines in other cells

polyimides and pentide therapy. The

22. The following are the names of the people who were present at the meeting:



Db	1907	QSDCLLSVTMP	SMVRHSLQ	TMLSVGYR	NIYTP	PDSS	TSFTIQ	DYSR	DGRJL	QTLHL	GTR	1966
Qy	2009	RVYKYGKLSK	LAETLYD	TTKVSFTY	DETAG	MLKTIN	QNEG	FTCT	TIRY	QIGPL	IDRQI	2068
Db	1967	RVLYKYTKQ	ARLSEV	LYDTTQ	VTLT	YEESG	VIKTIH	LMHDG	FICT	TIRY	QIGPL	IGRQI 2026
Qy	2069	FRFTEEGM	VNARFD	YNDNS	FRVT	SMQAV	INET	PLPID	LYRDD	VSGK	TEQFG	KFGVIYY 2128
Db	2027	FRFSEBGL	VNARFD	YSY-NN	FRVT	SMQAV	INET	PLPID	LYRYV	DVSG	RTEQFG	KFSVINY 2085
Qy	2129	DINQIIT	TAVMT	HTKH	FDAY	GRMKE	VQYEI	FRSL	MYMT	VQYD	NMGR	2175
Db	2086	DLNQVIT	TTVMK	HTKIF	SANG	QVIE	VQYEIL	KAIAY	WMTI	QYD	NVGR	2132

Search completed: June 24, 2004, 16:19:25  
Job time : 105 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:15:58 ; Search time 33 Seconds  
(without alignments)  
8071.343 Million cell updates/sec

Title: US-10-029-020-14  
Perfect score: 14887  
Sequence: 1 MDVKERKPYRSLTRRRDAER.....ELSDSANNIHFMRQSEMGR 2769

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14269	95.8	2825	2 T14271	Doc4 protein, stre
2	3804.5	25.6	2515	2 S47008	tenascin-like prot
3	3515.5	23.6	2406	2 A54148	odz protein - frui
4	3111.5	20.9	849	2 T46253	hypothetical prote
5	2109	14.2	2531	2 T16743	hypothetical prote
6	1175.5	7.9	782	2 A61625	tenascin-like prot
7	943	6.3	184	2 T12457	hypothetical prote
8	659.5	4.4	2019	1 JQ1322	tenascin precursor
9	652	4.4	1746	1 S19694	tenascin precursor
10	649.5	4.4	2201	2 A32160	tenascin-C - human
11	641	4.3	1810	1 A32230	tenascin precursor
12	626	4.2	4006	2 T09070	probable tenascin
13	592	4.0	3566	1 A40701	tenascin-X precurs
14	577.5	3.9	647	2 A43902	tenascin - eastern
15	572.5	3.8	4135	2 T42629	tenascin-X - bovin
16	438.5	2.9	2703	1 A24420	notch protein - fr
17	429	2.9	2524	2 A35844	Xotch protein - Af
18	428.5	2.9	2555	2 A40043	notch protein homo
19	427.5	2.9	2531	2 A46019	notch-1 protein -
20	424	2.8	1203	2 A49175	Notch B protein -
21	423.5	2.8	1064	2 A40136	fibropellin Ia - s
22	422.5	2.8	2531	2 S18188	notch protein homo
23	422	2.8	1620	2 T27283	hypothetical prote
24	420	2.8	2471	2 A49128	cell-fate determin
25	419.5	2.8	1220	2 A56136	jagged protein pre
26	413	2.8	2437	2 S42612	transmembrane prot
27	410	2.8	3191	2 T22945	hypothetical prote
28	400.5	2.7	1295	2 A32901	gpi protein precu
29	399	2.7	2531	2 T31070	notch homolog - se

30 392.5 2.6 1353 1 JH0675 restrictin precurs  
31 392.5 2.6 2321 2 S78549 notch3 protein - h  
32 388 2.6 1574 2 T13954 MEGF6 protein - ra  
33 387.5 2.6 1722 2 E89753 protein FllC7.4 [i  
34 387.5 2.6 1964 2 T09059 notch4 - mouse  
35 385.5 2.6 1356 2 A45445 janusin precursor,  
36 384 2.6 2318 2 S45306 notch 3 protein -  
37 383 2.6 1408 2 S16148 gene serrate prote  
38 382.5 2.6 2139 2 A35672 crumbs protein - f  
39 382 2.6 2352 2 T30201 Notch homolog prot  
40 381 2.6 861 2 A48825 Notch homolog Motc  
41 373.5 2.5 833 2 S19087 gene Delta protein  
42 372.5 2.5 832 2 A31246 neurogenic protein  
43 372.5 2.5 880 2 S00670 neurogenic repetit  
44 369.5 2.5 1111 2 T26972 hypothetical prote  
45 356.5 2.4 473 2 A56175 adhesive plaque pr

ALIGNMENTS

RESULT 1

T14271

Doc4 protein, stress-induced - mouse

N;Alternate names: odz protein homolog

C;Species: Mus musculus (house mouse)

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C;Accession: T14271

R;Wang, X.Z.; Kuroda, M.; Sok, J.; Batchvarova, N.; Kimmel, R.; Chung, P.; Zinszner, H.

EMBO J. 17, 3619-3630, 1998

A;Title: Identification of novel stress-induced genes downstream of chop.

A;Reference number: Z17951; MUID:98315054; PMID:9649432

A;Accession: T14271

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 1-2825 <WAN>

A;Cross-references: EMBL:AF059485; NID:g3170614; PID:g3170615; PIDN:AAC31807.1

C;Genetics:

A;Gene: Doc4

Query Match 95.8%; Score 14269; DB 2; Length 2825;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 2654; Conservative 47; Mismatches 52; Indels 88; Gaps 4;

Qy 1 MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
Db 1 MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
Qy 61 KDIVPQEAEEFCRTGANFTLRELGLLEEVTPPHGTLRYTDIGLPQCQGYSMGAGSDADMEAD 120  
Db 61 KDMVPQEAEEFCRTGTGNTLRELGLGEMTPPHGTLRYTDIGLPQCQGYSMGAGSDADLEAD 120  
Qy 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTET-----DHPGGL 170  
Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETGAPLHCSSASSTPIEQ 180  
Qy 165 -----DHPGGL 170  
Db 181 SPSPPPPPANESQRRLLGNGVAQPTPDSDEEEFVENSFLVKSGSASLGVAANDHPSSL 240  
Qy 171 QNHARLRTPPPPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQ 230  
Db 241 QNHPRLRTPPPPLPHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQ 300  
Qy 231 EPAHAQENWLLNSNIPLETNRNLGKQPFGLTLQDNLIEMDILGASRRHDGAYSDGHFLFKPG 290  
Db 301 EPTHAQDNWVLNSKIPVETRNLGKQPFGLTGWQDNLIEMDIFSASRRDGYSDGHFFFKPG 360  
Qy 291 GTSPLFCTTSPGYPLTSSVTYSPPPRPLPRSTFARPAFNLKPKSKYCNWKKAALSAIVIS 350  
Db 361 GTSPLFCTTSPGYPLTSSVTYSPPPRPLPRSTFARPAFNLKPKSKYCNWKKAALSAIVIS 420  
Qy 351 ATLVIILAYFVAMHLFGLNWLQPMEG--QMYEITDITASWPVPTDVSILYPSGGTGLT 408

Db 421 ATLVIILAYFVAMHLFGLNWHLQPMEGQMOMEIITEDTASSWPVPTDVSLYPSGGTGLT 480  
QY 409 PDRKGGKTTEGKPSFFPEDSFIDSEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNV 468  
Db 481 PDRKGGAAEGKPSLFPEDSFIDSEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNV 540  
QY 469 SLGKAALVGIYGRKGLPPSHTQFDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHET 528  
Db 541 SLGKAALVGIYGRKGLPPSHTQLDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHET 600  
QY 529 GFIOYLDGSIWHLAFYNDGKESEVVSFLTTAIESVDNCPNCGNGDCCISGTCCHFLGFL 588  
Db 601 GFIOYLDGSIWHLAFYNDGKESEVVSFLTTAIESVDNCPNCGNGDCCISGTCCHFLGFL 660  
QY 589 GPDGRASCPVLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTCIC 648  
Db 661 GPDGRASCPVLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTCIC 720  
QY 649 NPGYKGESCEBVCMDPTCSGRGVGVGECHCFVGGTNCETPRATCLDQCSHGHTFLP 708  
Db 721 NPGYKGESCEBVCMDPTCSSRGVGVGECHCFVGGTNCETPRATCLDQCSHGHTFLP 780  
QY 709 DTGLCSCDPSWTGHDCSIIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPCAEBHG 768  
Db 781 DTGLCNCDPSTGHDCSIIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPCAEBHG 840  
QY 769 TCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHCVQCLGWRGA 828  
Db 841 TCRDGKCECTPGWNGEHCTI-----EGCPGLCNGNGRCTLDLNGWHCVQCLGWRGT 891  
QY 829 GCDTSMETACGDSKNDGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQ 888  
Db 892 GCDTSMETCGDGDNDGDGLVDCMDPDCCLOPLCHVNPCLGSPDPLDIIQETQAPVSQ 951  
QY 889 QNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGNISFVNPL 948  
Db 952 QNLNSFYDRIKFLVGRDSTHISIPGENPFDGGHACVIRGQVMTSDGTPLVGNISFINNPL 1011  
QY 949 FGYTISRQDGSFDLVNTNGGISILRPERAPFITQEHTLWLPWDRFFVFMETIIVRHEENEI 1008  
Db 1012 FGYTISRQDGSFDLVNTNGGISILRPERAPFITQEHTLWLPWDRFFVFMETIIVRHEENEI 1071  
QY 1009 PSCDLNFPARNPVVSPSPLTSPASSCAEKGPVPEIQALQEEIISIGCKMRLSYLSSRT 1068  
Db 1072 PSRDLSNFPARNPVVSPSPLTSPASSCAEKGPVPEIQALQEEIIVGIAGCKMRLSYLSSRT 1131  
QY 1069 PGYKSVLRISLTHPTIPFNLKMKVHLMVAVEGRLFRKWFAPAAADLSYFIWDKTDVYNQKV 1128  
Db 1132 PGYKSVVRISLTHPTIPFNLKMKVHLMVAVEGRLFRKWFAPAAADLSYFIWDKTDVYNQKV 1191  
QY 1129 FGLSEAFVSVGYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHHALNIQSILHK 1188  
Db 1192 FGLSEAFVSVGYESCPDLILWEKRTAVLQGYEIDASKLGGWSLDKHHALNIQSILHK 1251  
QY 1189 GNGENOFVSQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLVVGDF 1248  
Db 1252 GNGENOFVSQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLVVGDF 1311  
QY 1249 NYIRRIFFPSGNVTNILELRNKRDFRHSHPAHKYIYLATDPMMSGAVFLSDSNRRVFKIKST 1308  
Db 1312 NYIRRIFFPSGNVTNILEM-----SHSPAHHYIYLATDPMMSGAVFLSDSNRRVFKIKST 1364  
QY 1309 VVVKDLVKNSEVWAGTGDCQCLPDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTM 1368  
Db 1365 TVVKDLVKNSEVWAGTGDCQCLPDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTM 1424  
QY 1369 IRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNV 1428  
Db 1425 IRRVDQNGIISTLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDNNV 1484  
QY 1429 VLOISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDE 1488

Db 1485 VLOISENHQVRIVAGRPMHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDE 1544  
QY 1489 KKINRIRQVTTSGEISLVAGAPSCDCCKNDANCDCFCGDDGYAKDAKLNTPSSSLAVCADG 1548  
Db 1545 KKINRIRQVTTSGEISLVAGAPSCDCCKNDANCDCFCGDDGYAKDAKLNTPSSSLAVCADG 1604  
QY 1549 ELYVADLGNIRIRFIRKKNKPFNTQNMVELSSPIDQELYLFTDTGKHLYTQSLPTGDYLY 1608  
Db 1605 ELYVADLGNIRIRFIRKKNKPFNTQNMVELSSPIDQELYLFTDTGKHLYTQSLPTGDYLY 1664  
QY 1609 NFTYTGDGDIITLITDNNGMVNVRRDSTGMPLWLVPDQGVYVWMTGNSALKSVTTQGH 1668  
Db 1665 NFTYTGDGDIITLITDNNGMVNVRRDSTGMPLWLVPDQGVYVWMTGNSALKSVTTQGH 1724  
QY 1669 ELAMTYHGNSGLLATKSNENGWTTFFYEYDSFGRLTNVTPTGQVSSFRSDTDSVHVQV 1728  
Db 1725 ELAMTYHGNSGLLATKSNENGWTTFFYEYDSFGRLTNVTPTGQVSSFRSDTDSVHVQV 1784  
QY 1729 ETSSKDDVTITTNLSASGAFYLLQDQVRNSYYIGADSLRLLLLANGMEVALQTEPHLLA 1788  
Db 1785 ETSSKDDVTITTHLSGSGAFYLLQDQVRNSYYIGADSLRLLLLANGMEVALQTEPHLLA 1844  
QY 1789 GTVNPTVGRNVTLPIIDNGLNLVWVRQKEQARGQVTVFGRRLRVHNRNLLSLDFRVT 1848  
Db 1845 GTVNPTVGRNVTLPIIDNGLNLVWVRQKEQARGQVTVFGRRLRVHNRNLLSLDFRVT 1904  
QY 1849 TEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVITYSPGGYIAGIQIGIMSERMEYD 1908  
Db 1905 TEKIYDDHRKFTLRILYDQAGRPSFWSPSSRLNGVNVITYSPGGHIAGIQIGIMSERMEYD 1964  
QY 1909 QAGRITSRIIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLSSVTMPNVARQTLET 1968  
Db 1965 QAGRITSRIIFADGKWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLSSVTMPNVARQTLET 2024  
QY 1969 IRSVGYVRNIYQOPPEGNASVIQDFTEDGHLHLLHTFYLGTGRRVIYKYGKLSKLAETLYDTT 2028  
Db 2025 IRSVGYVRNIYQOPPEGNASVIQDFTEDGHLHLLHTFYLGTGRRVIYKYGKLSKLAETLYDTT 2084  
QY 2029 KVSFTYDETAGMLKKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNAREFDNYDNS 2088  
Db 2085 KVSFTYDETAGMLKKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNAREFDNYDNS 2144  
QY 2089 FRVTSQCAVINETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINQIITTAVMTHSKHFDAY 2148  
Db 2145 FRVTSQCAVINETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINQIITTAVMTHSKHFDAY 2204  
QY 2149 GRMKEVQYEIIFRSLMYMTVQYDNMGRVVKKELKVGPYANTTRYSEYDADGQLQTVSIN 2208  
Db 2205 GRMKEVQYETFRSLMYMTVQYDNMGRVVKKELKVGPYANTTRYSEYDADGQLQTVSIN 2264  
QY 2209 DKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDDVQYKMDDEGFLRQGGDI 2268  
Db 2265 DKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDDVQYKMDDEGFLRQGGDV 2324  
QY 2269 FEYNSAGLLIKAYNRAGSWRYRYDGLGRRVSSKSSHHLQFFYADLTNPVKVTHLYN 2328  
Db 2325 FEYNSAGLLIKAYNRAGSWRYRYDGLGRRVSSKSSHHLQFFYADLTNPVKVTHLYS 2384  
QY 2329 HSSSITSLYDLOQHFLFAMELSSGDEFYIACDNICTPLAVFSGTGLMIKQIITYAYGEI 2388  
Db 2385 HSSSITSLYDLOQHFLFAMELSSGDEFYIACDNICTPLAVFSGTGLMIKQIITYAYGEI 2444  
QY 2389 YMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHMLWKHLSSSNVMPFN 2448  
Db 2445 YMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHMLWKHLSSSNVMPFH 2504  
QY 2449 LYMFKNNNPISNSQDIKCFMTDVNSWLLTFGQFQLHNVIPGYKPKDMDAMEPSYELIHTQM 2508  
Db 2505 LYMFKNNNPISNSQDIKCFMTDVNSWLLTFGQFQLHNVIPGYKPKDMDAMEPSYELVHTQM 2564  
QY 2509 KTOEWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKKFASSGSGVFGK 2568  
Db 2565 KTOEWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKKFASSGSGIFGK 2624



QY	1903	ERMEYDQAGRITSRIFADGKTWSYTYLEK--SMVLLHLSORQ--YIFEFDKNDRLSSVTM	1958
Db	1642	EAYSFDRNGRLNEIKYGDGSTMVYAFKDMFGSLPLKVTTPRSYLLQYDDAGALQSLTT	1701
QY	1959	P--NVARQTLTETIRSVGYRNIIYQPEGNASVIQDFTEDGHLLHTFYLTGTRRRVIKYGK	2016
Db	1702	PRGHIHAFSLQ--SLGFFKYQYSPINRHPPEILYNDEGQILAKIHPHQSGKVAFVHDT	1759
QY	2017	LSKLAETLYDTTKVSFTYDETAGMLKTLINLQNEGFTCTIPYRQIGPLIDRQIFRF--TEEG	2075
Db	1760	AGRLETILACLSSTHYTYQDTTSLVKSVEVQEPGFELRRBFKYHAGILKDEKLRFGSKNS	1819
QY	2076	MVNARFDVNDNSFRVTSMAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDI-----N	2131
Db	1820	LASARYKYAYDGNARLSGIEMAIDDKELPTTRYKYSONLQGLE-----VVQDLKITRN	1872
QY	2132	QIITTAVMTHTKHF-----DAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVVKELKVGP	2185
Db	1873	AFNRTVIQDSAKQFAIVDYDQHGVRKSVLMNVKNIDVPRLELDYDLNRRIKSQKTTFG-	1931
QY	2186	YANTTRYSEYDADGQ--LQTVSINDKPLWRYSYDLNGN--LHLLSPGNSARLTPRYDIRD	2243
Db	1932	-RSTAFDKINYNADGHVVEVLGTNN---WKYLFDENGNTVGVDVQGEKFN---LGYDIGD	1984
QY	2244	RITRLGDVQY--KMEDEGFLRQRGGDIFEYNSAGLLIKAYNRA--GSWSVRYRYDGLGRRV	2300
Db	1985	RVIKVGDFVFNNDYDARGFVVKRGEQKYRNNRQGLIHSFERERFQSW---YYYDDRSRLV	2041
QY	2301	SSKSSHSHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQGLHFLAMELSSGDEFYIAC	2360
Db	2042	AWHDNKGNNTQYYYANPRTPLVTHVHPFKISRTMKLFYDDRDMLIALE--HEDQRYVAT	2100
QY	2361	DNIGTPLAVFSGTGLMIKQILLYTAYGEIYMDTNPNFQIILGYHGLYDPLTKLVHMGRRD	2420
Db	2101	DQNGSPLAFFDQNGSIVKEMKRPFPGRIIKDTKPEFFVPIDFHGGLIDPHTKLIVYTEQRQ	2160
QY	2421	YDVLAGRWTSPDHELWKHLSSSNVMPFN--LYMFKNNPNISNSQDIKCFMTDVNSWLLTF	2478
Db	2161	YDPHVQWMTF--LWETLATENSHPTDVFYIYRHNNDP--NPKNPQNYMIDLSWLQLF	2216
QY	2479	GFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTOEWDNSKS-----ILGVQCEVQKQLK	2531
Db	2217	GYDLNN-----MQSSRYTKLAQYTPQASIKSNTLAPDFGVISGLECIVEK---2261	
QY	2532	AFVTLERPDQLYGSTITSCQAPKTKKFASSGS---VFGKGVKFALKDGRVTTDIISVA	2587
Db	2262	---TSEKPSDFDFVPKPLLTKEPKVRNLLPRVSYRRGVFEGVLLSRIGGRALVSVDGS	2318
QY	2588	N---EDGRRVAAILNHAHYLENLHFTIDGVDTHYFVRPGPSEGLAILGLSGRRRTLENG	2644
Db	2319	NSVVQD--VVSSVFNNSYFLD-LHFSIHDQDVIFYVK---DNVLKLRDDNEELRRLGGM	2371
QY	2645	VNVTVSQINTVLNGRTRRYTDIQLQYGCALCLNTRYGTTLDEEKARVLELARQAVRQAWA	2704
Db	2372	FNISTHEISDHGGSAK---EURLHGPDAVVIKYGVDPQEQRHRIILKHAHRAVERAWE	2428
QY	2705	REQQRLREGEGLRAWTEGEKQVLTSTGRVQYDGFVVISVEQYPELSDSANNIHFMRQS	2764
Db	2429	LEKQLVAAGFQGRGDWTEEEKELVQHGDVDGWNGIDIHSHKYPQLADDPGNVAFQRDA	2488
QY	2765	EMGR 2769	
Db	2489	KRKRR 2493	
RESULT 3			
A54148			
cdz protein - fruit fly (Drosophila sp.)			
C:Species: Drosophila sp.			
C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 21-Jul-2003			
C:Accession: A54148			
R:Levine, A.; Bashan-Ahrend, A.; Budai-Hadrian, O.; Gartenberg, D.; Menasherow, S.; Wide			

Cell 77, 587-598, 1994			
A;Title: odd Oz: a novel Drosophila pair rule gene.			
A;Reference number: A54148; MUID:94243925; PMID:7514504			
A;Accession: A54148			
A;Status: preliminary; not compared with conceptual translation			
A;Molecule type: mRNA			
A;Residues: 1-2406 <LEV>			
C;Genetics:			
A;Gene: FlyBase:Ten-m			
A;Cross-references: FlyBase:FBgn0004449			
F;514-540/Domain: EGF homology <EGF>			
F;610-637/Domain: EGF homology <EGF1>			
Query Match 23.6%; Score 3515.5; DB 2; Length 2406;			
Best Local Similarity 34.0%; Pred. No. 5.7e-187;			
Matches 724; Conservative 433; Mismatches 843; Indels 130; Gaps 42;			
QY	384	EDTASWPVPTDVSLYPSGGTGLETPDRKGKTTEGKPSFFPPEDSFIDSGEIDVGERRAS	443
Db	347	EATSSAATSSSQSLTPSLSSSLANANNGGARTF---PARSFPPDG-TTFGQITLQOKLT	402
QY	444	QKIPPGTFWRSQVFIHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQDFVELLDGRRLL	503
Db	403	KEIQPYSYWNMQFYQSEPAYVKFDYTIIPRGASIGVYGRNALPHTQYHFKEVLGFSFAS	462
QY	504	TOEARS--LEGTPROSRGTVPSSSHETGFIQYLDSGIWHLAFYN--DGKESEVVSFLTAT	560
Db	463	TRTARAAHLSITREVT-----YMEPGHWFVSLYNDGDDVDVQELTFYAAVAE	508
QY	561	ESVNDNCPNCGNGDCISGTCHCFGLGPDCCGRASCVPVLCSGNGQYMKGRCLCHSGWK	620
Db	509	DMTQNCNPGCSGNGQCLLGHCCNCPGFGDDCSESVCPVLCSEQHGEYTNCEICNPGWK	568
QY	621	AECDVPTNQCIDVACSNHGTCTGTCTICNPGYKGESCEBEVD CMDPTCSGRGVCVRGECH	680
Db	569	KECSLRHDECEVADCSGHGHCVSQKQCMRGYKGFCEVD CPHNCSGHGFCADGTIC	628
QY	681	FVGWGTNCET--PRATCLDQCSGHGTFLPDTGLCSDPSPSWTGHDCSIEICAAACGGHG	737
Db	629	KKGWKGPDCATMDQDALQCLPDCSGHGTFLDTQTCTCEAKWSGDDCSKELCDLDCGQHG	688
QY	738	VCVGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHVLDRVVK	797
Db	689	RCEGDACADPEWGGECYCNTELCDVRCNEHGGQCKGTCLCVTGWNGKHCTI-----	739
QY	798	EGCPGLCNGNGRCTLNLNG--WHCVCOLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPD	856
Db	740	EGCNSCAGHQCRVSGEGQWECECYEGWDGPDCCGIALELNCGDSKNDKDKGLVDCEDPE	799
QY	857	CCLQLCHINPLCLGSPNPLDIQETQVPVSQQNLHSPFYDRKFLVGRDSTHIIIPGENPF	916
Db	800	CCASHVCKTSQLCVSAPKPIVDLLRKQPPAITA---SFFERMKFLIDESLQNYAKLET	856
QY	917	DGHGACVIRGQVMTSDGTPLVGVNISFVNNPLFGYTIISQDGSFDLVTNGGISILRPER	976
Db	857	NESRAVIRGRVVTSLGMLGVGRVS--TTTLLEGFTLTRDDGWFDLMVNGGGAVTLQFGR	915
QY	977	APFTTQEHTLWLPWDRFFVNMETIIMRHEENE-----IPSCDLSNFARPNVVPSPPLTS	1030
Db	916	APFREQSRIVQVPWNEVVIIDLVMMSMSEKGLAVTTHTTCTFAHDYDLMKPVVLLASWKHG	975
QY	1031	FASSCAEKGPVPEIQALQBEHISISGCKMRLSYLSRTTPGYKSVLRISLTHPTIPFNLMK	1090
Db	976	FOGACPDRAISAILAESQVIESLQIPGTGLNLVHSSRAAGYLSTIKLQLTDPDVIPTSLHL	1035
QY	1091	VHLMVAVEGRLEFRKFAAAPDLSYFYIWDKTDVYNQKVFGLSEAFVSUVGYEYSCPDLL	1150
Db	1036	IHLRITIEGILFERIFEADPGIKFTYAWNRLNIYRQRYGVTTAVVKVGYQYTDCTD-IV	1094
QY	1151	WEXRTTVLQGYEIDASKLGSWLDKHHALNIQSGLHKGNGENQFVSQQPPVIGSIMNG	1210
Db	1095	WDIQTTLKSGHMSISEVGVGNWLDIHHRYNPFHEGILQKGDGNSIYLRNKPRIILTMGDG	1154







Db 417 HVSLLAAEEAEGVSTTCPNDCSGRGSCVLGKDCIDGYQGVDCKSVCPVLCSAHGYGG 476  
QY 610 GRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTITGTCICNPGYKGESCEEVDCMDPTCSG 669  
Db 477 GVCHCEEGWKGAECDDIPVGECEVPNCSSHGRCIEGEGCHCERGWKGPYCDQHDCDPLCSG 536  
QY 670 RGVCVRGECHCFVGGWGTNCET---PRATCLDQCSCGHGTFLDDTGLCSCDPSWTGHDCSI 726  
Db 537 HGTCVAGQCYCKAGWQGEDCGTIDQVYQCLPGCSEHGTYDLETGQCVCERHWTGPDSCQ 596  
QY 727 EICAAADCGGHGVCGGTCTCRCEDEGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHC 786  
Db 597 AVCSLDCGRNGVCESGKCRNSGWTGNLQDLPDSDRCSRHGQCKNGTVCVCSQGWNGRHC 656  
QY 787 TIAHYLDRVVKEGCPGLCNGNGRCRTLDLNGWHVCQLGWRGAGCDTSMETACGDSKDNHG 846  
Db 657 TL-----PGCENGCSRHGQCTLENGEYRCDCEGWAGSDCSIALELNCKNDINDG 707  
QY 847 DGLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVVSQQNLHSFYDRIKFL 901  
Db 708 DGMTDCSDSECCSHPACSEHIMCLSSNDPVEVLLRKQPP---SVTASFYQVRVKFL 759

RESULT 7  
T12457  
hypothetical protein DKFZp564O0423.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 23-Jul-1999  
R:Accession: T12457  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z17524  
A:Accession: T12457  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-184 <WAM>  
A:Cross-references: EMBL:AL080120  
A:Experimental source: fetal brain; clone DKFZp564O0423  
C:Genetics:  
A>Note: DKFZp564O0423.1

Query Match 6.3%; Score 943; DB 2; Length 184;  
Best Local Similarity 99.5%; Pred. No. 1.8e-45;  
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 2587 ANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGDAILGLSGGRRITLNGVN 2646  
Db 2 ANEDGRRVAAILDHAHYLENLHFTIDGVDTHYFVKPGPSEGDAILGLSGGRRITLNGVN 61  
QY 2647 VTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDDEEKARVLELARQVRQAWARE 2706  
Db 62 VTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDDEEKARVLELARQVRQAWARE 121  
QY 2707 QORLREGHEGLRAWTEGEKQQLVLTGRVQYDGFVVISVEQYPELSDSANNIHFMRQSEM 2766  
Db 122 QORLREGHEGLRAWTEGEKQQLVLTGRVQYDGFVVISVEQYPELSDSANNIHFMRQSEM 181  
QY 2767 GRR 2769  
Db 182 GRR 184

RESULT 8  
JQ1322  
tenascin precursor - mouse  
N:Alternate names: contactin; hexabrachion  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000.  
C:Accession: JQ1322; A37936; B37936; S14571; S50209  
R:Saga, Y.; Tsukamoto, T.; Jing, N.; Kusakabe, M.; Sakakura, T.  
Gene 104, 177-185, 1991  
A:Title: Murine tenascin: cDNA cloning, structure and temporal expression of isoforms.

A:Reference number: JQ1322; MUID:92009211; PMID:1717349  
A:Accession: JQ1322  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-2019 <SAG>  
A:Cross-references: GB:D90343; NID:9220609; PIDN:BAA14355.1; PID:9220610  
A:Experimental source: cell line 2H6GR  
A>Note: the authors translated the codon ATG for residue 60 as Trp  
R:Weller, A.; Beck, S.; Ekblom, P.  
J. Cell Biol. 112, 355-362, 1991  
A:Title: Amino acid sequence of mouse tenascin and differential expression of two tenascin  
A:Reference number: A37936; MUID:91107734; PMID:1703162  
A:Accession: A37936  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 1026-  
A:Cross-references: GB:X56304  
A:Accession: B37936  
A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 1026-  
A:Cross-references: GB:X56304  
R:Weller, A.; Beck, S.; Ekblom, P.  
submitted to the EMBL Data Library, August 1990  
A:Description: Amino acid sequence of mouse tenascin and differential expression of two t  
A:Reference number: S14571  
A:Accession: S14571  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-201, 'E', 203-317, 'S', 319-1018, 'S', 1020-1024, 'H', 1026-1305, 'S', 1307-2019 <WE  
A:Cross-references: EMBL:X56304; NID:954768; PIDN:CAA39751.1; PID:954769  
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.  
Biochim. Biophys. Acta 1219, 613-622, 1994  
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by  
A:Reference number: S50209; MUID:95035091; PMID:7524681  
A:Accession: S50209  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 46-146 <GLU>  
A:Cross-references: EMBL:X80281  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type  
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracell  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-2019/Product: tenascin, long splice form #status predicted <MAT>  
F:23-1071,1527-2019/Product: tenascin, short splice form #status predicted <MAT2>  
F:408-434/Domain: EGF homology <EGF>  
F:622-703/Domain: fibronectin type III repeat homology <FN3A>  
F:711-793/Domain: fibronectin type III repeat homology <FN3B>  
F:802-884/Domain: fibronectin type III repeat homology <FN3C>  
F:892-976/Domain: fibronectin type III repeat homology <FN3D>  
F:984-1064/Domain: fibronectin type III repeat homology <FN3E>  
F:1073-1155/Domain: fibronectin type III repeat homology <FN3F>  
F:1346-1428/Domain: fibronectin type III repeat homology <FN3G>  
F:1437-1519/Domain: fibronectin type III repeat homology <FN3H>  
F:1527-1608/Domain: fibronectin type III repeat homology <FN3I>  
F:1616-1696/Domain: fibronectin type III repeat homology <FN3J>  
F:1704-1784/Domain: fibronectin type III repeat homology <FN3K>  
F:1799-2007/Domain: fibrinogen beta/gamma homology <FEG>  
F:38,166,184,327,788,1018,1079,1093,1119,1184,1210,1275,1301,1354,1364,1394,1627,1878,19

Query Match 4.4%; Score 659.5; DB 1; Length 2019;  
Best Local Similarity 19.5%; Pred. No. 5.4e-28;  
Matches 407; Conservative 273; Mismatches 714; Indels 689; Gaps 97;  
QY 566 CPSNCYGNGDCISGTCFCLGFLGPDCCGRASCPVLCSGNGQYMKGRCLCHSGWKGAEC-- 623  
Db 315 CPNDCFDRGRCTGTCYCEGFTGEDCGELTCPNDCQGRGQCEGQCVCNEGAGADCSE 374  
QY 624 -----DVPTNQCIDVACSNHGTCTITGTCICNPGYKGES 656  
Db 375 KRCPADCHHRCLNGQCECDGDTGADCGDLQCPN-GCSGHRGVNGQCVCDGEGYTGED 433

QY	657	CEEVD	CMDPTCSGRGVCVRGECHCFVGGWGTNCETPRATCLDQC	SGHGF	FLPDTGL	CSCD	716																																		
Db	434	CSORR	CPN-DCHNRGLCVQ	GKCI	CEQGFK	GDCS--EMSCPNDCHQHGR	CV--NGMCICD 488																																		
QY	717	PSWTH	DCSIEICAA	DCGGHVCVGGTCR	CECDGWMGAACDQ	RACHPRCAEHG	TCRDGKCE 776																																		
Db	489	DDYTG	EDCDRDRR	CDSPRDCSQRGR	CVGQCICEDG	TGPDCAELSCPSDCHGHGR	CVNGQCI 548																																		
QY	777	CSPGW	NGEHC	TI	AHYLD	RVVKEGCPGLCNGNGRC	TDLNGWHCV	COLGWRGAG	CDTSMET 836																																
Db	549	CHEG	TGKDC	-----EQR	CPSDCHGQGR	-----EDGQC	ICHEG	FTGLDCG	---QR 592																																
QY	837	ACGDS	KONDG	DGLV	CDMPDC-CLQ	PLCHINPLCLUGSPNPLD	IIQETQ	VPVSQ	QNLSFY 895																																
Db	593	SCPND	CSNQ	-----QCV	SGRCICNEG	YTID--CSEVSP	PKDLI--VTEV	TEETV	NLAW 643																																
QY	896	DR---	IKFLVGR	DSTHI-----IP	GENPFDG	HACVIRGQVMTSDG	TPLVGN	---	940																																
Db	644	DNEMR	VTLEYL	IMYTP	THADGLE	MQFRVPGD-----QT	STTIRE	LEPGVE	YFI 690																																
QY	941	--ISF	VNPLFGY	TISRQD	GSFDLV	TNGGISIILP	PERAPFITQ	EHTLW	LPWDRFFV	MET 998																															
Db	691	RVFAI	LENKR-S	IPVSAR	VATYLP	PAPEG---LK	FSIK-ETS	VEVEW	DPLDIA	ETWE 743																															
QY	999	IIMR	HEENEI	PCDLS	NFARNP	VVSPSPLT	SFASSCA	EKGPI	VP	PEIQALOE	SISGCK 1058																														
Db	744	IIFRN	NKDEGE	ITKSL	RRPE-----TSY	QTGLAPG-----Q	EY	IS	---	782																															
QY	1059	MRLSY	LSRTPG	YKSVLR	ISLTH	PTIPFNLM	KVHLMV--AVEGR	LFRK	WFAA	APDLSY	YF 1116																														
Db	783	LHIV	KONT	RGPLK	KVTT	TRLDAPS-----H	IEVKD	VTDT	TALIT	WFKPLA	EID--- 831																														
QY	1117	IWDK	TDVYNQ	KVFL	SEAFV	SVGYEYES	CPDLIL	WEKRT	TVLQ	GEY	EIDASKLG	WSLDKH 1176																													
Db	832	-----	SI	ELSYGI	KDVP	G-----DRT	TIDL	THED	NQYS	IGNLR	PDTE 868																														
QY	1177	HALNI	QSGIL	HKNG	ENQFVS	QPPVIGS	IMNGRR	RSI--SCP	SCNGL	ADGNK	LAPVAL 1235																														
Db	869	YEV	SLIS-----	RRV	MASN	PAKET	FTITG--LD	AP---	896																																
QY	1236	TCGSD	GSLYV	GFNY	IRRI	FP	SGNVTNILEL	RN	KDFR	SHSPA	KYYLATD	PMGSAV	FLS 1295																												
Db	897	-----	ENL	RRVSQ	TDNSITL-----E	WRNV	KADID	SYRI	KYAP	ISGG--- 933																															
QY	1296	DSNRR	VFKIK	STVV	VKDLV	KNSEV	VAGTG	DQC	LPDD	TRCGD	GKGATE	ATL	NP	GITV 1355																											
Db	934	-----	DH	AEID	VPKS	QQA-----	TTK	TTL	GLR	P	GT- 959																														
QY	1356	DKFGL	IYF	VDG	T	MIRRID	QNGIIST	L	GSND	LTSAR	PLSCDS	VMDIS	QVR	LE--WPT	DLA 1413																										
Db	960	-EYGI	-----GVS	AVKGD	KESDP	ATINA	ATEID	APKDL--RV	SETT	QD	SLTFF	WTPLA 1010																													
QY	1414	-----IN	-----PMD	NSLY	VLDN	NVVLQ	ISENHQ	VRIV	AGR	PMHC----- 1448																															
Db	1011	KFDRY	RLNYS	LPTG	QSM	EVQLPK	DATSH	VLTD--LE	PGQ	EYTV	LLIAE	KGRHKS	KPARV 1067																												
QY	1449	-----QV	PGIDH	FLLS	SKVA-----	I	HAT	LES	A	TALAV	SHN---GVLY 1482																														
Db	1068	KASTE	EVPS	LENL	TVTE	AGWDGL	RLNWT	ADDL	AYE	YFV	IQVE	ANNV	ETA	HNFTV	PGNLR 1127																										
QY	1483	IAET	DEK	KIN	RIRQ	VTT	SGEIS	LV	AGAP	SGCD	CKND	ANC	DC	FC	SGDDGY-----AK	DAKL 1536																									
Db	1128	AADIP	GLK-----V	ATSV	RV	SIY	GVA-----R	G	YRT	PVLS	AE	T	STG 1163																												
QY	1537	NTPSS	LAV	CAD	GEL	YVAD	LGNIR	IRFIR	KNK	PF	LNTQ	NMYEL	SS	PIDQ	E	LYL	FD	TG	KHL 1596																						
Db	1164	TTPNL	-----G	EV	TV	AEV	G-----W	D	AL---	T 1182																															
QY	1597	Y	TQ	S	L	P	T	G	D	I	T	L	T	D	N	G	N	M	V	N	V	R	D	S	T	G	M	P	L	W	L	V	P	D	G	-----	1647				
Db	1183	LNW	T	A	P	E	G	A	Y	K	N	F	I	Q	V	L	E	A	D	T	T	Q	T	V	Q	N	-----	L	T	V	P	G	L	R	S	V	D	L	P	G	L 1225
QY	1648	---QVY	WVTM-----GT	NS	A	L	K	S	V	T	T	Q	G	E	L	A	M	T	Y	H	G	N	S	G	L	L	A	T	K	S	N	E	N	G	-----W	T	T 1693				

Db	1226	KAATRYIILRGVTQDFGT--APLSVEVLTEDLPOL-----GGLSVTEVSWDGLTLNWTT	1278
QY	1694	FYEYDSFGLTNVTFPTGOVSSFRSDTDSVVHVQVETSSKDDVTITTNLSASGAFYTL--	1751
Db	1279	-----DDLAYKHVVVQVQVEANNVEAAQNLTVPGSLRAVDI	1313
QY	1752	----LQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRNVTLLPIDNG	1807
Db	1314	PGLKADTPYRVSIYGVIOGYRTPMLSTDVSTARE-----PEIGNLVSDVTPKS	1362
QY	1808	LNLVEWRQRKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQ	1867
Db	1363	FNL-SW-----TATDGI--MFTIEII--	1382
QY	1868	AGRPSSLWSPSRLNGVNVITYSPGG-----YIAGIQRGIMSERM-----EYDQA	1910
Db	1383	DSNRLLOQTAEHNISAERTAHISGLPPSTDFIVYLSGIAPSIRTKTITSTATTEAEPEVD	1442
QY	1911	GRITSRIFADGKTWSYTYLE--KSMVLLLSQRQYIFEFDKNDRLLSSVTMPNVAR-QTL	1966
Db	1443	NULLVS DATPDGPRLSWTADGIFDSFVIRDTK-----KQSEPOEISLPSPERTDI	1495
QY	1967	ETIRSVGYR-NIYQPPEGNASVIQDFTEDGHLLHTFYLTGTRRVIYKYKLSKLAETL-	2024
Db	1496	TGLREATYEIELYGISRGRS-----QPVSAIATTAMGSPKEIM--FSDITENAATVS	1547
QY	2025	--YDTTKVS---FTY-DETAGMLKTINLQNEGFTCTIRYRQIGPLIDBQIF-----Rfte	2073
Db	1548	WRAPTAQVESPRITYVPMTCGAPSMVTV--DGTDTETRLVKLTGPGVEIRVSVIAMKGFE	1605
QY	2074	EGMVNARFDYNDNS-----FRVTSMQAVINETP--LPIDLYRD-----DVSG	2115
Db	1606	SDPVSGLTITALDCPSGLLIANITDSEALAMWQPAIATVDSYVISYTGervpEVRTVSG	1665
QY	2116	KTEQFGKFGV-----IYDINQIITTAVMT-HTKHFDAYGRMKVEQYEIFRSLMYW	2165
Db	1666	NTVEYELHDLBATEYILSIFAEGQOKSSTIATKFTTDLDSPREFTATEVQSETALLTW	1725
QY	2166	MTVQYDNMGRVVKELKVGPPYANTTRYSYBYDA-DGQLQTVSINDKPLWRYSYDLGNLH	2224
Db	1726	RP-----PRASVTGYLLVYESVDGTVKEVIGPDTT---SYSLAD---	1762
QY	2225	LLSPGN--SARLTPLRYDIRD1-----TRLGDVQYKMDDEGFLRQRGGDIFEYNSAGL-	2276
Db	1763	-LSPSTHYSARIQALSGLSRSLKIQTIFTIG-LLYPFPRDCSQAMLNGD---TTSGLY	1816
QY	2277	-----LIKAY---NRAGWSVRVRYDGLGRRVSSKSSHHLQF-----FYAD	2316
Db	1817	TIYINGDKTOALEVYCDMTSDGGGWIVFLRRKN-GREDFYRNWKAYAAAGFDRREEFWLG	1875
QY	2317	LTNPTKVTHLVNHSSEBITSLYYDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLM	2376
Db	1876	LDNLSKITAQGY-----ELRVDLQDH-----GESAYAVYDRFSVG-----	1911
QY	2377	IKQILYTAIGEYMDTNPNFQI-IIGYHGLYDPLTKLVHMGR	2418
Db	1912	-----DAKSRYKLKVEGYSGTAGDSMN--YHNGR	1938

RESULT 9  
S19694

tenascin precursor - pig

N;Alternate names: contactin; hexabrachion

C/Species: *Sus scrofa domestica* (domestic pig)

```
C:\pepper> dir c:\corpora\compendex (compressed) p-97
```

C>Date:	10-Sep-1999	#sequence	revision	10-Sep-1999	#text	change	10-Sep-1999
C:\pepper>							

C;Accession: S19694

R;Nishi, T.; Weinstein, J.; Gillespie, W.M.; Paulson, J.C.

Eur. J. Biochem. 202, 643-648, 1991

Eur: U. Biochem. 202; 643-646, 1971  
 A: Title: Complete primary structure of porcine tenascin. Detection of tenascin transcripts

A:Reference number: S19694; MUID:92104189; PMID:1722152

A:Accession: S196694

A: Accession: SI3034  
A: Molecule type: mRNA



A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-2201 <RES>  
A;Cross-references: EMBL:X78565; NID:G556844; PIDN:CAA55309.1; PID:G556845  
R;Gulcher, J.R.; Nies, D.E.; Marton, L.S.; Stefansson, K.  
Proc. Natl. Acad. Sci. U.S.A. 86, 1588-1592, 1989  
A;Title: An alternatively spliced region of the human hexabrachion contains a repeat of  
A;Reference number: A32160; MUID:89160821; PMID:2466295  
A;Accession: A32160  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 431-538, 'R', 540-1065, 'H', 1067-1599, 'LWLHPRASN', 1609-2054, 'L', <GUL>  
A;Cross-references: GB:M24630; NID:G514363; PIDN:AAA52703.1; PID:G553348  
R;Siri, A.; Carnemolla, B.; Saginati, M.; Lepini, A.; Casari, G.; Baralle, F.; Zardi, L  
Nucleic Acids Res. 19, 525-531, 1991  
A;Title: Human tenascin: primary structure, pre-mRNA splicing patterns and localization  
A;Reference number: S14015; MUID:91187670; PMID:1707164  
A;Accession: S14015  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-243, 245-369, 'V', 371-679, 'R', 681-1676, 'I', 1678-2139, 'TR', 2142, 2144-2201 <SI  
A;Cross-references: EMBL:X56160; NID:G37226; PIDN:CAA39628.1; PID:G37227  
R;Nies, D.E.; Hemesath, T.J.; Kim, J.H.; Gulcher, J.R.; Stefansson, K.  
J. Biol. Chem. 266, 2818-2823, 1991  
A;Title: The complete cDNA sequence of human hexabrachion (tenascin). A multidomain prob  
A;Reference number: S16166; MUID:91131572; PMID:1704365  
A;Accession: S16166  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-538, 'R', 540-1065, 'H', 1067-1599, 'LWLHPRASN', 1609-2054, 'LH', 2055-2201 <NIE>  
A;Cross-references: EMBL:M55618; NID:G184483; PIDN:AAA88083.1; PID:G184484  
R;Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.  
Biochim. Biophys. Acta 1219, 613-622, 1994  
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by  
A;Reference number: S50206; MUID:95035091; PMID:7524681  
A;Accession: S50206  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 46-125 <GLU>  
A;Cross-references: EMBL:X80280  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C;Genetics:  
A;Gene: GDB:HXB  
A;Cross-references: GDB:120073; OMIM:187380  
A;Map position: 9q33-9q33  
C;Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type  
C;Keywords: alternative splicing; extracellular matrix  
F;408-434/Domain: EGF homology <EGF>  
F;622-703/Domain: fibronectin type III repeat homology <FN3A>  
F;711-794/Domain: fibronectin type III repeat homology <FN3B>  
F;892-976/Domain: fibronectin type III repeat homology <FN3C>  
F;1437-1519/Domain: fibronectin type III repeat homology <FN3D>  
F;1619-1790/Domain: fibronectin type III repeat homology <FN3E>  
F;1709-1790/Domain: fibronectin type III repeat homology <FN3F>  
F;1798-1878/Domain: fibronectin type III repeat homology <FN3G>  
F;1886-1966/Domain: fibronectin type III repeat homology <FN3H>  
F;1981-2189/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 4.4%; Score 649.5; DB 2; Length 2201;

Best Local Similarity 22.3%; Pred. No. 2.2e-27;

Matches 310; Conservative 166; Mismatches 492; Indels 423; Gaps 64;

QY 545 NDGKESEVSVFLTTAIESVDNCPNSCYNGDCTISGTCFGLGFLGPCGRASCPVLCS-G 603

Db 208 DDGFTGEDCSQLA-----CPSDCNDQKCVNGVCICFEGYAGADCSREICPVCSEE 259

QY 604 NGQYMKGRCLCHSGWGAECDDVPTNQCIDVACSNHGTCITGTICINPGYKGESCEEVDCM 663

Db 260 HGTCVDGLCVCHDGFAGDDCNKPL--CLN-NCVNRGRCVNEECVCDGFTGEDCSELICP 316

QY 564 DPTCSGRGVCVRGECHCFVWGWTNCETPRATCLDQCSGHGTFPLPTGLCSCDSPSWTGH 723

Db 317 N-DCFDRGRCTGTCYCEEGFTGEDCGKP--TCPHACTQGR--EEGQCVCDEGFAGLD 371

QY 724 CSIEICAAADCGHGVCVGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSPGMWG 783  
Db 372 CSEKRCPADCHNRGRCDGRCECDDGFTGADCGELKCPNGCSGHGRCVNGCVCDEGYTG 431  
QY 784 EHCT-----IAHYLDRVVKEG-----CPGLCNGNGRCTLDLNGWHCV 820  
Db 432 EDCSQLRCPNDCHSRGCV-EGKCVCEQKGYGDCSDMSCPNDCHQHGR-----VNGM-CV 486  
QY 821 COLGWRGAGC-----DTSMETACGDSKONDGDLV--DCMDPDC-----CL--Q 860  
Db 487 CDDGTGEDCDRDRQCPRDCSNRGLCVDGQCVCEDEGTGPDCAELSCPNDCHGQRCVNGQ 546  
QY 861 PLCH--INPLCLGSPNPLDIQETQVPVSQQNLHLSFYDRIKFLVGRDSTHIIPGENPFD 917  
Db 547 CVCHGEFMGKDCKEQRCPSDCHGQRCVDGQCICHEGF-----TGLDC----- 589  
QY 918 GGHAC-----VIRGOVMTSDGTPLVGVNISFVNNPLFGYTISRQDGSFDLVTNGGI 968  
Db 590 GQHSCPSDCNNLGQCVSGRCICNEG--YSGEDCSEVSPK-----DLV----- 630  
QY 969 SIILRFERAPFITQ--EHTLWLPWD-RFFVVMETIIMRHEENEIPSCDLSNFARPNPVWSP 1025  
Db 631 -----VTEVTEETVNLANDNEMRVTEYLV-----VYTP 658  
QY 1026 SPLTSFASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISL---THP 1082  
Db 659 T-----HEGGLMQFRVPGDQTS-----TIIQELEPGVEYFIRVFAILENKK 700  
QY 1083 TTPFNLMKVHLMVAVEGRLEFRKWFAPAAADLSYYFTWDKTDV-----YNQKVF- 1130  
Db 701 SIPVSARVATYLPAPPEGLKFK-----SIKETSVEVEWDPLDIAFETWEIIFRNMNKEDEGE 756  
QY 1131 -----LSEAFVSVCYEYESCPDLI-----LWEKRTTVLQG---YEI---D 1164  
Db 757 ITKSLRRRPETSYRQTGLAPCQEYEISLHVKNNTRGPLKRVTTTRLDAPSQIEVKDVT 816  
QY 1165 ASKLGW--SLDKXHALNTQSGI-----LHXNGENQFVSQQPPVIGSIMNGRRR 1213  
Db 817 TTALITWFKPLAEIDGIELTYGIKDVPGDRTTIDLTEDENY-----SIGNLXPD-TEY 869  
QY 1214 SISCPCSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRFPFSGNVTNILELRNKDFRH 1273  
Db 870 EVSLISRRGDMSSNP--AKETFTTGLDAP-----RNLRRVSQTDNSITL-----EWRN 915  
QY 1274 SHSPAKEYLATDPMGAVFLSDSNSRRRVFKIKSVVVKDLVKNSEVAGTGDQCLPFDD 1333  
Db 916 GKAAIDSYRIKYAPISGG-----DHAEDVDVPKSQLA----- 946  
QY 1334 TRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIISTLLGSNDLTSARPL 1393  
Db 947 -----TTKTTLTGLEPGT--EYGI-----GVSARKEDKESNPATINAAATELTPKDL 991  
QY 1394 SCDSVMDISQVRLEWPTDLA-----IN-----PMDNSLYVLDNNVVLIQISE 1434  
Db 992 QVSETAETS-LTLLWKTFPLAKFDRLNYSLPTGQWVGVLPRNTTSYVLRG---LEPGQ 1047  
QY 1435 NHQVRIVAGRPMHC-----QVPGIDHFLLSKVAIHATLESATAL--AVSHNGVL 1481  
Db 1048 EYNVLLTAEKGRHKSKPARVKASTEQAPELENLTIVTEVGWDGLRLNWTAAOQAYEH---- 1103  
QY 1482 YIAETDE-KKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDFSGDDGYAKDAKLNTPS 1540  
Db 1104 FIIQVQEANKVEAARNLTVP--SLRAVDIPLGKAATPYTVSIYGIQY-----RTPV 1155  
QY 1541 SLAVCADGELYVADLGNIRIRFIRKNKPFNLNTQNMYELSSPIDQELYLFDTTGKHLYTQS 1600  
Db 1156 LSAAEASTGE--TPNLGEVVVAEVGWDALKLN-----WT 1186  
QY 1601 LPTGDYLYNFTYTGDDITLITDNNNGNMVNVRDRDSTGMPLWLVPDQVYVWVTMGTSAL 1660  
Db 1187 APEGAYEYFFIQVQEADTVAAQN-----LTVPGLRSTDLPLGLKAAT 1229



Db 920 -TTRATLTGLRP-----GTEYGIGVTAVRQ-DRESAPATINAGTDLDPKDLVSDPTE 971

QY 1401 ISQVRLEWPTDLA-----INPMDNSLYVLDDNNVVLQISENHQVRIV 1441

Db 972 -TTLSLRWRRPVAKFDRYRLTYVSPSGKKNEMEIPVDSTSFILRG---LDAGTEYTISLV 1027

QY 1442 AGRPMHCQVP-----GIDHFLLSKVAIHATLESATALAVSHNGV 1480

Db 1028 AEKGRHKSPTTIKGSTEEPELGNLSVSETGWGDFQLTWTADGAYE-----NFV 1078

QY 1481 LYIAETDEKKINRIRQVTTSGEISLVAGAPS-----GCDCKNDANCDCFSGDDGYAKDAKL 1536

Db 1079 IQVQOSDNPE-----ETWNITVPGGQHSVNVTVGLKANTPYNVTLYGVIRGYRTKPLY 1130

QY 1537 NTPSSLAVCADGELYVADL-----GNIRIRFIRKNK---PF-----L 1570

Db 1131 VETTIGAHPVEVGELTVSDITPESFNLSWTTTNGDFDAFTIEIIDSNRLLPEMFEFNISGNS 1190

QY 1571 NTQNMVELSSPIDQELYLFDTTGKHLX-----TQSLPTGDYLYNFTYTTGDGDITL 1620

Db 1191 RTAHISGLSPSTDIFIVLYGIS--HGFRTOAISAAAATTEAEPEVDNLLVSDATPDGFRLT 1248

QY 1621 ITDNNG-----NMVNVVRDSTGMPLWLVP-----DGQVYWVTMGT 1656

Db 1249 WTADDGVFDSFVLKIRDTKRKSD--PLELIVPGHERTHDITGLKEGTEYEIELYGVSSGR 1306

QY 1657 NS-ALKSVTTQGHELAMMTYHGN-SGLLATKSNEN---GWTTFYEYDSFGRLTNVTFPT 1710

Db 1307 RSQPINSVAT-----TWVGSPKGISFSDITENSARVSWTP-----PR 1343

QY 1711 GQVSSFRSDTSSVHVQVETSSKDDVTITTNLSASGAFYTLQDQVRNSYYIGADGSLRL 1770

Db 1344 SRVDSYRVS-----YVPITGGTPNVVTVDGSKTR-----TKLVKLVP-----GVDYNVNI 1388

QY 1771 LLANGMEVALQTEPHLLAGTVNPTVGKRN--VTLPIDNGLNLVEWR-----QRK 1817

Db 1389 ISVKGFEE--ESEP--ISGILKTALDSPSGLVVMNITDSEALATWQPAIAAVDNYIVSYS 1443

QY 1818 EQARGQVT--VFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTRLRLYDQA--GRPSL 1873

Db 1444 SEDEPEVTQMVSGNTVE-YDLNGLRPATEYTLRVHAVKDAQKSETLSQTFTGLDAPKDL 1502

QY 1874 WSPSSRLNGVNVITYSPGYIAGIQRGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSM 1933

Db 1503 SATEVQSETAVITWRPP-----RAPVTDYLLTYESIDGRVKEVILDDETTSYTLTE--- 1553

QY 1934 VLLIHSQRQYIFEFDKNDRLSSVTMPNVARQILETIRSVGYRYNIQOPPEGNASVIOQDFT 1993

Db 1554 ---LSPSTQYTVKLQALSR----SMRSKMIQIVFTTIGLLY----PYPKDCSQALLNGEV 1602

QY 1994 EDGHLHLHTFYLGTR 2008

Db 1603 TSG--LYTIYLANGDR 1615

RESULT 12

T09070

probable tenascin X - mouse

C;Species: Mus musculus (house mouse)

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jan-2000

C;Accession: T09070

R;Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; So

A;Description: Sequence of the mouse major histocompatibility locus class III region.

A;Reference number: Z16543

A;Accession: T09070

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-4006 <ROW>

A;Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564958

C;Genetics:

A;Gene: TNX

A;Map position: 17

A;Introns: 124/1; 735/1; 773/3; 826/1; 914/1; 1037/1; 1135/1; 1232/1; 1329/1; 1440/1; 1

O19/1; 3115/1; 3208/1; 3302/1; 3405/1; 3517/1; 3558/1; 3606/1; 3646/1; 3694/1; 3737/3;

C;Superfamily: tenascin-X; EGF homology; fibrinogen beta/gamma homology; fibronectin ty

C;keywords: extracellular matrix

F;422-448/Domain: EGF homology <EGF>

F;826-906/Domain: fibronectin type III repeat homology <3FR>

F;3789-3997/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 4.2%; Score 626; DB 2; Length 4006;

Best Local Similarity 22.4%; Pred. No. 1.2e-25;

Matches 359; Conservative 163; Mismatches 585; Indels 498; Gaps 77;

QY 566 CPNCYGNGDCISGTCCHCFLGLGPDCCGRASCPVLCSNGQYMKGRCLCHSGWKGAECDV 625

Db 422 CPRDCRGRCEDEGVVCVCHAGYSGEDCGVRSCTPGDCRGRNCESGRVCVWPGYTGRDCG- 480

QY 626 PTNOCIDVACSNHGTCTGTCTICNPYKGESCEEVDCTPDCSGRGVCVRGEGCHCFVGMG 685

Db 481 -TRAC-PGDCRGRGRVDCRVCNPGFTGEDCGSRR-C-PGDCRGHGHCENGVCVCAVGYS 537

QY 686 GTNCTPRATCLDQCSGHGTFLPDTGLCSDPSPWTHDCSIEICAADCGGHGVCVGGTCR 745

Db 538 GDCST--RSCPSDCRGRGQCL--NGLCECEGYSGEDCGIRRCPRDCSQHGVCQDGLCM 593

QY 746 CEDGWMGAACDQACHPRCAEHGTCRDGKCECSPGNGEHCTIAHYLDRVVKEGCPGLCN 805

Db 594 CHAGYAGEDCSIRTCPADCRRRGRCEDEGRCVCNPGYGPAC-----ATRITCPADCR 644

QY 806 GNGRCTLDLNGWHCVQLGWRGAGCDTSMETA-----CGDSKNDGDGLVDCMD---P 855

Db 645 GRGRVCQGV---CMCYVGYSGEDCGQEPPASACPGGCG-PRELCRAGQCVCVEGFRGP 699

QY 856 DCCUQPL--CHINPLCL-----GSG 872

Db 700 DCAITQCPGDCRSRGECIQGRVCVCEGYAGDDCGEEIPAIQNMRMHLLLETTVREWTRA 759

QY 873 PNPDIIEQETQVPVSQQNLHSPYDRIKFLVGR-DSTHILPGENPFDDGGHACVIRGQVMTS 931

Db 760 PGVDAYEIQFIPMTGVSPPFTARVPSSASAYDQRLAPGD-----YQVTVRALRGTS 814

QY 932 DGTPLVGVNISFVNNPLFGYISRODGSFDLVNTNGGISILRFERAFFITQEHTLWLWP- 990

Db 815 WGPFAASKTITTMIDGP-----QDLRVAVT-----PTTLDLSWL 848

QY 991 -----DRFFVMETIIMRHEEN-EIPSCDLSNFARNPNVVPSPPLTSF----- 1031

Db 849 RPOAEVDRFVVSYSAGNQVRLEVP-----PEADRTQLTDLMPGVEYVVTVA 897

QY 1032 -----ASSCAEKGPVPEIQALQEEIISIGCKMRLSYLSSRTPGYKSVLRISLTHP 1082

Db 898 ERGHAVSYPASIRANTGSLPSGLLEATDEPPPSG-----PSTTQGAQAPILILEHH- 948

QY 1083 TTPFNLMKVHLMVAVEGRLEFRKWFAAAPDLSYF-----IWDXTDV----- 1123

Db 949 --PLGELKV-LGRDKAGRLSVAV-TAQPDSTFAHFQLRMQVAEGPWAHEELLPGDVQQALV 1004

QY 1124 -----YNQKVFGLSE-AFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLD 1174

Db 1005 PPPPPGAPYKLFHLHGITPGGKISVPITYQGIMD-----RAEQPGKPSVQPRLGLTVT 1058

QY 1175 KHALNIQSGILHKNGENQFVS-----QQPPVICSMGNRRRSISPCSCNGLADGN 1227

Db 1059 ---GLTSDSLLHWTVPGEPSDFLIQYKDKDGPQAISVEGQRPSTPIS-----GLEPGR 1110

QY 1228 K-----LLAPVALTCGSD-----GSLYVGD-----FNVIIR 1253

Db 1111 KYKFILYGLIGKKRHGPLMAEAKILIQSDPDGSPPLRIGELWVTDPTPHSLHLSWTVLGG 1170

QY 1254 IFPSGNVTNILELRNKDFR-----HSHSPAHHYY-----LA 1284

Db 1171 QFDS----FVQYRDKEGQPRVVPVEGPDRSVVISLDPNRKYRFTLFGIANKKRYGPLT 1226

QY 1285 TDPMSGAVFLSDSNS-RRVFKIKSTVVVKDLVKNSEVVA-GTGD-----QCL 1329

Db 1227 ADGTTAAETKESSEPPRLGELTVTGTPDLSRLSWTVARGPFDSEFVILYKDAQGQPSV 1286  
QY 1330 PFDDTRCGDGGKATEATLTNPRGITVDKFLIYFVDGTMIRRIDQNGIISTLLGSNDLTS 1389  
Db 1287 PIE-----GDENEVTPGLESNRKYKMNLYGL-----RGRQVGPVSVV-----AKT 1328  
QY 1390 ARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLD---NNVVLQ--ISENH-----QV 1438  
Db 1329 APPDVVDQPLSPPEPSTEAP-EPPEKPLRLRELIVTDTTPNSLSLSWTVSEGQDFQFVIQY 1387  
QY 1439 RIVAGRPMHCQVPG-----IDH-FLLSKVAIHATLESATALAVSHNGVLYIAET 1486  
Db 1388 KNGDGGQPKVVRVPGHEDQVTISGLEPDHKKYKMNLYGIH-----SGQRVGPFVSVGVTA 1441  
QY 1487 DEKKINRIRQVITS-----GEISLVAGAPSGCCDKNDANCDFS-----GD-DG 1529  
Db 1442 EEPPSPTEPTEAPEPTPEEPLLGELTVTGSSP-----DSLSTWTVPPQGFDS 1490  
QY 1530 YAKDAKLNTPSSLAVCADGELYVADLG-----NIRIRFKNKPFLNTQNMVELS-----S 1580  
Db 1491 FTVQYK-----SSDGRPQVMRVGGQEREIVRDLDPGHKY--KMNLYGLHGHVRVG 1539  
QY 1581 PIDQELYLFDITGKHLTYQSLPTGDLNFTYTGDDITLTNDNNGMNVNRRDSTGMPL 1640  
Db 1540 PVS-----TVGTTDKQ-PEEKPPLEPRLEELTVTG-----VTPNS-----VSL 1576  
QY 1641 WLVPDQG-----VYVVTMGNTSALKSVTTQGHELAM-----MTYHGNSG----- 1680  
Db 1577 SWIVPEGQFDSFVQYQDRDQGPREVPLTAEQREVTVLDDPARIYKMNLYGLHGHVRVG 1636  
QY 1681 -----LLATKSNE-----NGWTTFYEY--DSFGRLTNVTFPTGOVSSFR-----SD 1719  
Db 1637 PLSVSAMTAPAPATEASKLLPEPRLGELTVTDVTPDSVGLLW--TVPEGEFDSFTVQYKD 1694  
QY 1720 TDSSVHVQVETSSKDDVTITTNLSAGAFYTLT---QDQVRNSYVIGADGSLRLLLANGM 1776  
Db 1695 RDGQPHVSVVAADLREVTI-PGLEPSSKYKFLFVGQDGKRSQV-----SV 1740  
QY 1777 EVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEWR-----ORKEQARGQVT 1825  
Db 1741 EAKTATHDDTSPG-APPRLGELWVTDFTDLSRL-SWTVPEGHFDSFVQFKDKSGPRVV 1798  
QY 1826 VEGRRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDAGR 1870  
Db 1799 PVKQGER-----TATITLADAGRKYRF-LLYGLLGR 1828

RESULT 13

A40701  
tenascin-X precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Dec-1999  
C;Accession: A40701; A33725; C42175  
R;Bristow, J.; Tee, M.K.; Gitelman, S.E.; Mellon, S.H.; Miller, W.L.  
J. Cell Biol. 122, 265-278, 1993  
A;Title: Tenascin-X: a novel extracellular matrix protein encoded by the human XB gene  
A;Reference number: A40701; MUID:93300909; PMID:7686164  
A;Accession: A40701  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-3566 <BRI>  
A;Cross-references: EMBL:X71937  
R;Morel, Y.; Bristow, J.; Gitelman, S.E.; Miller, W.L.  
Proc. Natl. Acad. Sci. U.S.A. 86, 6582-6586, 1989  
A;Title: Transcript encoded on the opposite strand of the human steroid 21-hydroxylase/c  
A;Reference number: A33725; MUID:89367293; PMID:2475872  
A;Accession: A33725  
A;Molecule type: mRNA  
A;Residues: 2748-3199, 'V', 3201-3298, 'E', 3299-3314, 'G', 3316-3566 <MOR>  
A;Cross-references: GB:M25813; NID:g183069; PIDN:AAA35884.1; PID:g183070  
R;Matsumoto, K.; Arai, M.; Ishihara, N.; Ando, A.; Inoko, H.; Ikemura, T.  
Genomics 12, 485-491, 1992

A;Title: Cluster of fibronectin type III repeats found in the human major histocompatibi  
enascin.  
A;Reference number: A42175; MUID:92217969; PMID:1373119  
A;Accession: C42175  
A;Molecule type: DNA  
A;Residues: 1849-1936 <MAT>  
A;Experimental source: clone 3.9kf3-1  
A;Note: sequence extracted from NCBI backbone (NCBIP:95694)  
C;Genetics:  
A;Gene: GDB:TNXA; D6S103E; TNX; XA; XB  
A;Cross-references: GDB:568487; OMIM:600261  
A;Map position: 6p21.3-6p21.3  
C;Superfamily: tenascin-X; EGF homology; fibrinogen beta/gamma homology; fibronectin typ  
C;Keywords: extracellular matrix; glycoprotein  
F;435-461/Domain: EGF homology <EGF>  
F;748-828/Domain: fibronectin type III repeat homology <3F1>  
F;829-856/Domain: fibronectin type III repeat homology #status atypical <3F2>  
F;873-953/Domain: fibronectin type III repeat homology <3F3>  
F;975-1055/Domain: fibronectin type III repeat homology <3F4>  
F;1078-1158/Domain: fibronectin type III repeat homology <3F5>  
F;1167-1247/Domain: fibronectin type III repeat homology <3F6>  
F;1248-1317/Domain: fibronectin type III repeat homology #status atypical <3F7>  
F;1323-1403/Domain: fibronectin type III repeat homology <3F8>  
F;1412-1492/Domain: fibronectin type III repeat homology <3F9>  
F;1510-1590/Domain: fibronectin type III repeat homology <3F10>  
F;1618-1749/Domain: fibronectin type III repeat homology #status atypical <3F11>  
F;1678-1749/Domain: fibronectin type III repeat homology <3F13>  
F;1751-1831/Domain: fibronectin type III repeat homology <3F14>  
F;1849-1929/Domain: fibronectin type III repeat homology <3F15>  
F;1955-2035/Domain: fibronectin type III repeat homology <3F16>  
F;2061-2141/Domain: fibronectin type III repeat homology <3F17>  
F;2167-2246/Domain: fibronectin type III repeat homology <3F18>  
F;2274-2354/Domain: fibronectin type III repeat homology <3F19>  
F;2382-2462/Domain: fibronectin type III repeat homology <3F20>  
F;2488-2568/Domain: fibronectin type III repeat homology <3F21>  
F;2584-2664/Domain: fibronectin type III repeat homology <3F22>  
F;2677-2757/Domain: fibronectin type III repeat homology <3F23>  
F;2771-2851/Domain: fibronectin type III repeat homology <3F24>  
F;2878-2958/Domain: fibronectin type III repeat homology <3F25>  
F;2977-3067/Domain: fibronectin type III repeat homology <3F26>  
F;3078-3159/Domain: fibronectin type III repeat homology <3F27>  
F;3167-3247/Domain: fibronectin type III repeat homology <3F28>  
F;3255-3334/Domain: fibronectin type III repeat homology <FBG>  
F;3349-3557/Domain: fibrinogen beta/gamma homology <FBG>  
  
Query Match 4.0%; Score 592; DB 1; Length 3566;  
Best Local Similarity 21.4%; Pred.No. 7.8e-24;  
Matches 364; Conservative 182; Mismatches 549; Indels 606; Gaps 86;  
  
QY 562 SVDNCPNVCYNGDNCISGTCCHFLGFLGPDCCGRASCPVLCSGNGQYMKGRCLCHSGWKGA 621  
Db 369 STRTCPRDCRGRCEDEGEICDGTGSGDDCGVRSCTPGDCNQRCEDGRCVCPWPGYGT 428  
QY 622 ECDVPTNQCIDVACSNHGTCTITGTCICNPGYKGESCEVEVDCMDPTCSGRGVCVRGECHCF 681  
Db 429 DCG---SRACPRDCRGRGRCENGVVCNAGYSGEDCGVRSC-PGDCRGRGRCSGRCMCW 484  
QY 682 VWGWTNCHT-----PRATCLD-----QCSGHGTFLLPDTGL 712  
Db 485 PGYTGRDCGTRACPGDCRGRGRVCVDRGVCPNPGFTGEDCGSRRCPCGDCRHHG--LCEDGV 542  
QY 713 CSCDPSWTGHDCSIEICAAACGGHGVVGVGTCTCEDGWMGAACDQACHPRCAEHGTCRD 772  
Db 543 CVCDAGYSGEDCSTRSCPGGCRGRGQCLDGRVCEDGYSGEDCGVRQCPNDCSQHGVQCQD 602  
QY 773 GKCECSPGNNGEHCCTIAHYLDRVVVKEGCPGLCNGNGRCTLDLNGWHVCVQLGWRGAGCDT 832  
Db 603 GVCICWEGYVSEDCSI-----RTCPSNCHGRGR-----EEGRCLCDPGYTGTCAT 649  
QY 833 SMETACGSKDNDGDGLVDCMDPDCCLOP--LCHIN-----PLCLGSPNPLD 877  
Db 650 RMCPPA-----DCRGRGRVCVGVCLCHVGYGGEDCGQEPPASACPGGCGPRE 696





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:15:08 ; Search time 85 Seconds  
(without alignments)  
10278.469 Million cell updates/sec

Title: US-10-029-020-14  
Perfect score: 14887  
Sequence: 1 MDVKERKPYRSLTRRDAER.....ELSDSANNIHFVRSQSEMGR 2769

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL 25:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriaph:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14529	97.6	2771	11 Q9WTS7	Q9wts7 mus musculus
2	14269	95.8	2825	11 Q70465	Q70465 mus musculus
3	12033.5	80.8	2824	13 Q9W7R3	Q9w7r3 brachydanio
4	10227	68.7	2715	11 Q9WTS6	Q9wts6 mus musculus
5	9968	67.0	2192	13 Q804R1	Q804r1 brachydanio
6	9642	64.8	2705	13 Q9W6V6	Q9w6v6 gallus gall
7	9616.5	64.6	2802	13 Q9DER5	Q9der5 gallus gall
8	9610.5	64.6	2764	11 Q9WTS5	Q9wts5 mus musculus
9	9607	64.5	2731	11 Q9WTS4	Q9wts4 mus musculus
10	9601	64.5	2765	11 Q9R1K2	Q9r1k2 rattus norv
11	9501	63.8	2725	4 Q9UKZ4	Q9ukz4 homo sapien
12	9303	62.5	2590	13 Q9W7R4	Q9w7r4 brachydanio
13	9145	61.4	2346	11 Q9JLC1	Q9jlc1 mus musculus
14	8059.5	54.1	2144	4 Q9ULU2	Q9ulu2 homo sapien
15	7063	47.4	1828	11 Q80TD2	Q80td2 mus musculus
16	6774	45.5	1769	4 Q9P273	Q9p273 homo sapien

17	6189	41.6	1198	11 Q80TF5	Q80tf5 mus musculu
18	5638	37.9	1086	4 Q9P2P4	Q9p2p4 homo sapien
19	5614	37.7	1071	4 Q7Z3C7	Q7z3c7 homo sapien
20	3979.5	26.7	1045	4 Q9NVW1	Q9nvw1 homo sapien
21	3854	25.9	730	4 Q96MS6	Q96ms6 homo sapien
22	3822	25.7	2731	5 Q61307	Q61307 drosophila
23	3820.5	25.7	2731	5 Q9VNU6	Q9vnu6 drosophila
24	3813	25.6	2731	5 Q18366	Q18366 drosophila
25	3804.5	25.6	2515	5 Q24551	Q24551 drosophila
26	3753	25.2	991	11 Q8CAT1	Q8cat1 mus musculu
27	3730	25.1	964	4 Q9NV77	Q9nv77 homo sapien
28	3517	23.6	3004	5 Q24550	Q24550 drosophila
29	3515	23.6	3004	5 Q9VYN8	Q9vyn8 drosophila
30	3398.5	22.8	930	11 Q9JLC0	Q9jlc0 mus musculu
31	3111.5	20.9	849	4 Q9NT68	Q9nt68 homo sapien
32	2378	16.0	831	13 Q9PU49	Q9pu49 gallus gall
33	2351	15.8	625	4 Q96SY2	Q96sy2 homo sapien
34	2108.5	14.2	2560	5 Q21980	Q21980 caenorhabdi
35	2079.5	14.0	2531	5 Q8MPZ2	Q8mpz2 caenorhabdi
36	1464	9.8	337	11 Q9R1K0	Q9r1k0 rattus norv
37	1336	9.0	278	13 Q9DEQ8	Q9deg8 gallus gall
38	1244	8.4	442	4 Q9NZJ2	Q9nzj2 homo sapien
39	1129	7.6	536	11 Q8C8D2	Q8c8d2 mus musculu
40	1120.5	7.5	272	11 Q9R1J9	Q9r1j9 rattus norv
41	1093.5	7.3	266	11 Q9R1K1	Q9r1k1 rattus norv
42	1089	7.3	229	11 Q9QYZ1	Q9qyz1 mus musculu
43	1036	7.0	1009	5 Q8MQJ6	Q8mgj6 drosophila
44	1034.5	6.9	329	11 Q8BSL5	Q8bsl5 mus musculu
45	943	6.3	184	4 Q9Y4S2	Q9y4s2 homo sapien

ALIGNMENTS

RESULT 1  
Q9WTS7  
ID Q9WTS7 PRELIMINARY; PRT; 2771 AA.  
AC Q9WTS7-  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Ten-m4  
GN NRGI OR ODZ4 OR TEN-M4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Balb/c; TISSUE=Brain;  
RA Ohashi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,  
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;  
RT "Mouse Ten-m/Odz is a new family of dimeric type II transmembrane  
RT proteins expressed in many tissues.";  
RL J. Cell Biol. 0:0-0(1999).

Result No. Score Query Match Length DB ID Description

Query Match		97.6%;	Score 14529;	DB 11;	Length 2771;		
Best Local Similarity		97.0%;	Pred. No. 0;				
Matches 2688;		Conservative 40;	Mismatches 41;	Indels 2;	Gaps 1;		
QY	1	MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60				
Db	1	MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60				
QY	61	KDIVPQEAEEFCRTGANFTLRELGLBEETPPHGTLYRTDGLPQCYSMGAGSDADMEAD	120				
Db	61	KDMVPQEAEEFCRTGNTFTLRELGLGEMTPPHGTLYRTDGLPHCGYSMGASSDADLEAD	120				
QY	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEHENTETDHPGGLQNHARLRTPP	180				
Db	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEHENTETDHPSSLQNHPRLRTPP	180				
QY	181	PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSLSGEPAGGAQEPAHQAENWL	240				
Db	181	PPLPHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHLSLSGEPAGSAQEPTHAQDNWL	240				
QY	241	LNSNIPLERNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS	300				
Db	241	LNSNIPLERNLGKQPFGLTQDNLIEMDILSASRHDGAYSDGHFLFKPGGTSPLFCTTS	300				
QY	301	PGYPLTSSIVYSPPPRPLPRSTFEAPAFNLKKPSKYNKKAALSAIVISATLVILLAYF	360				
Db	301	PGYPLTSSIVYSPPPRPLPRSTFSRPAFNLKKPSKYNKKAALSAAILISATLVILLAYF	360				
QY	361	VAMHLFGLNWLQPMEG- -QMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTE	418				
Db	361	VAMHLFGLNWLQPMEGQMOMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGAEE	420				
QY	419	GKPSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFDHPVHLKFNVS LGKAALVGI	478				
Db	421	GKPSLFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFDHPVHLKFNVS LGKAALVGI	480				
QY	479	YGRKGLPPSHTQDFVELLDGRRLLTQEARSLEGTPRQSRGTVPSSSHETGFIQYLDSGI	538				
Db	481	YGRKGLPPSHTQDFVELLDGRRLLTQEARSLEGTPRQSRGVPVPPSSHETGFIQYLDSGI	540				
QY	539	WHLAFYNDGKESEWSFLTTAIESVDNCPNCGYNGDCISGTCHCFGLGFLGPDGCRASCP	598				
Db	541	WHLAFYNDGKESEWSFLTTAIESVDNCPNCGYNGDCISGTCHCFGLGFLGPDGCRASCP	600				
QY	599	VLCSENGQYMKGRCLCHSGWKGAECQVPTNQCIDVACSNHGTCTITGTCICNPGYKGESCE	658				
Db	601	VLCSENGQYMKGRCLCHSGWKGAECQVPTNQCIDVACSSHGTCTIMGTICINPGYKGESCE	660				
QY	659	EVDCMDPTCSRGVGVGECHCFVGGTNCETPRATCLDQCSGHGTFPLPDTGLCSCDPS	718				
Db	661	EVDCMDPTCSSRGVGVGECHCSVGWGTNCETPRATCLDQCSGHGTFPLPDTGLCNC DPS	720				
QY	719	WTGHDCSIEICAADCGGHGVGVGTCRCEDGWMGAACQDQACHPRCAEHGTCRDGKCECS	778				
Db	721	WTGHDCSIEICAADCGGHGVGVGTCRCEDGWMGAACQDQACHPRCAEHGTCRDGKCECS	780				
QY	779	PGWNGEHCITIAHYLDRVVKEGCPCGLCNGNGRCTLDLNGHCHVCQLGWRGAGCDTSMETAC	838				
Db	781	PGWNGEHCITIAHYLDRVVKEGCPCGLCNGNGRCTLDLNGHCHVCQLGWRGTGCDTSMETGC	840				
QY	839	GDSKNDGDGLVDCMDPCCCLQPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRI	898				
Db	841	GDGKNDGDGLVDCMDPCCCLQPLCHVNPCLGSPDPDLIIQETQAPVSQQNLNPFYDRI	900				
QY	899	KFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFVNNPLFGYTISRQDG	958				
Db	901	KFLVGRDSTHSIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFINNPLFGYTISRQDG	960				
QY	959	SFDLVNCGISIIILRFERAPFITQEHTLWLPWDRFFVNMETIIMRHEENEIIPSCDLSNFAR	1018				
Db	961	SFDLVNCGISIIILRFERAPFITQEHTLWLPWDRFFVNMETIVMRHEENEIIPSCDLSNFAR	1020				
QY	1019	PNPVVSPSPLTSFASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRIS	1078				

Db	1021	PNPVVSPSPLTSFASSCAEKGPIVPEIQALQEEIIVAGCKMRLSYLSSRTPGYKSVLRIS	1080				
QY	1079	LTHPTIPENLMKVHLMVAVEGRLEPRKWFAPDLSYFIWMDKTDVYNQKVFGLSEAFVS	1138				
Db	1081	LTHPTIPENLMKVHLMVAVEGRLEPRKWFAPDLSYFIWMDKTDVYNQKVFGEAFVS	1140				
QY	1139	GYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHHALNIQSGILHKGNGENQFVSQ	1198				
Db	1141	GYEYESCPDLILWEKRTAVLQGYEIDASKLGWSLDKHHALNIQSGILHKGNGENQFVSQ	1200				
QY	1199	QPPVIGSIMNGRRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPSG	1258				
Db	1201	QPPVIGSIMNGRRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPSG	1260				
QY	1259	NVTNILELRNKDFRHSHPAHKYIATDPMGAVFLSDNSRRRVFKIKSTVVVKDLVKNS	1318				
Db	1261	NVTNILEMRNKDFRHSHPAHKYIATDPMGAVFLSDNSRRRVFKVKSSTTVVKDLVKNS	1320				
QY	1319	EVAGTGDCQCLPFDDTRCGDGGKATEATLTNPRGITVDKFLIYFVDGTMIRRIDQNGII	1378				
Db	1321	EVAGTGDCQCLPFDDTRCGDGGKATEATLTNPRGITVDKFLIYFVDGTMIRRIDQNGII	1380				
QY	1379	STLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQV	1438				
Db	1381	STLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQV	1440				
QY	1439	RIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRQVT	1498				
Db	1441	RIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRQVT	1500				
QY	1499	TSGEISLVAGAPSGCDCKNDANCDGDDGYAKDAKLNTPSSLAVCADGELYVADLGN	1558				
Db	1501	TSGEISLVAGAPSGCDCKNDANCDGDDGYAKDAKLNTPSSLAVCADGELYVADLGN	1560				
QY	1559	RIRFIRKNKPFNTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDLYLNFTYTGDDI	1618				
Db	1561	RIRFIRKNKPFNTQNMVELSSPIDQELYLFDTSKHLYTQSLPTGDLYLNFTYTGDDI	1620				
QY	1619	TLITDNNGMVNVRRDSTGMPMLVVPDQVYVVTMTGNSALKSVTTOGHELANMTYHGN	1678				
Db	1621	THITDNNGMVNVRRDSTGMPMLVVPDQVYVVTMTGNSALRSVTTQGHELANMTYHGN	1680				
QY	1679	SGLLATKSNENGWTTFFEYDSFGRLTNTVFTPTGOVSSFRSDTSSVHVQVETSSKDDVTI	1738				
Db	1681	SGLLATKSNENGWTTFFEYDSFGRLTNTVFTPTGOVSSFRSDTSSVHVQVETSSKDDVTI	1740				
QY	1739	TTNLSASGAFYTLQDQVRNSYVIGADGSLRLLLLANGMEVALQTEPHLLAGTVNPTVGKR	1798				
Db	1741	TTNLSASGAFYTLQDQVRNSYVIGADGSLRLLLLANGMEVALQTEPHLLAGTVNPTVGKR	1800				
QY	1799	NVTLPIDNGLNLVEWRQKEQARGQVTFGRRLRVHNRNLLSLDFDRTREKTYDDHRK	1858				
Db	1801	NVTLPIDNGLNLVEWRQKEQARGQVTFGRRLRVHNRNLLSLDFDRTREKTYDDHRK	1860				
QY	1859	FTLRILYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQGIMSERMEYDQAGRITSRIF	1918				
Db	1861	FTLRILYDQAGRPSLWSPSSRLNGVNTYSPGGHIAGIQGIMSERMEYDQAGRITSRIF	1920				
QY	1919	ADGKTWSYTYLEKSMVLLHLSORQYIFEDKNDRLSSVTMPNVARQTLETIRSVGYRNI	1978				
Db	1921	ADGKWSYTYLEKSMVLLHLSORQYIFEDKNDRLSSVTMPNVARQTLETIRSVGYRNI	1980				
QY	1979	YQPPEGNASVIQDFTEDGHLHTFYLGTGRRVYKYGKLSKLAETLYDTTKVSFTYDETA	2038				
Db	1981	YQPPEGNASVIQDFTEDGHLHTFYLGTGRRVYKYGKLSKLAETLYDTTKVSFTYDETA	2040				
QY	2039	GMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGVNARFDYDNDNSFRVTSMQAVI	2098				
Db	2041	GMLKTIVNLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGVNARFDYDNDNSFRVTSMQAVI	2100				
QY	2099	NETPLPIDLYRDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHFDAYGRMKEVOYEI	2158				

Db 2101 NETPLPIDLYRDDVSGKTEQFGKFGVIYVDINQIITAVMTHTKHFDAYGRMKEVQYEI 2160

QY 2159 FRSLMYMTVQYDNMGRVVKELXGVPYANTTRYSYEYDADGQLQTVSINDKPLWRYSD 2218

Db 2161 FRSLMYMTVQYDNMGRVVKELXGVPYANTTRYSYEYDADGQLQTVSINDKPLWRYSD 2220

QY 2219 LINGNLHLLSPGNSARLTPLRYDIRDRIITRLGDVQYKMDDEGFLRQRGDIFEYNSAGLLI 2278

Db 2221 LINGNLHLLSPGNSARLTPLRYDLDRITRLGDVQYKMDDEGFLRQRGDVFYNSAGLLI 2280

QY 2279 KAYNRAGWSVRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSSEITSLY 2338

Db 2281 KAYNRAGWSVRYDGLGRRVSSKSSHSHLQFFYADLTNPTKVTHLYNHSSSEITSLY 2340

QY 2339 YDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILLYTAYGEIYMDTNPFI 2398

Db 2341 YDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILLYTAYGEIYMDTNPFI 2400

QY 2399 IIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHKLWHLSSNVMPFNLYMFKNNPI 2458

Db 2401 IIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHKLWHLSSNVIPFFHLYMFKNNPI 2460

QY 2459 SNSQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDSKS 2518

Db 2461 SNSQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELVHTQMKTQEWDSKS 2520

QY 2519 ILGVQCEVQKQKAFVTLERFDQLYGTITSCQAPKTKKFASSGSVFGKGVKFALKDGR 2578

Db 2521 ILGVQCEVQKQKAFVTLERFDQLYGTITSCQAPETKKFASSGSIFGKGVKFALKDGR 2580

QY 2579 VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGR 2638

Db 2581 VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGR 2640

QY 2639 RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEBKARVLELARQRA 2698

Db 2641 RTLENGVNVTVSQINTMLSGRTRRYTDIQLQYRALCLNTRYGTTVDEEKVRVLELARQRA 2700

QY 2699 VRQAWAREQORLREGEGRLAWTEGEKQQVLS\*GRVQGYDGFVVISVEQYPELSDSANNI 2758

Db 2701 VRQAWAREQORLREGEGRLAWTDGEKQQVLTGRVQGYDGFVTVSVEQYPELSDSANNI 2760

QY 2759 HFMRQSEMGR 2769

Db 2761 HFMRQSEMGR 2771

RESULT 2  
O70465

ID O70465 PRELIMINARY; PRT; 2825 AA.

AC O70465;

DT 01-AUG-1998 (TReMBLrel. 07, Created)

DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE DOC4.

GN NRG1 OR OD24 OR DOC4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NTH/Swiss;

RX MEDLINE=98315054; PubMed=9649432;

RA Wang X.Z., Kuroda M., Sok J., Batchvarova N., Kimmel R., Chung P.,

RA Zinszner H., Ron D.;

RT "Identification of novel stress-induced genes downstream of chop.;"

RL EMBO J. 17:3619-3630(1998).

DR EMBL; AF059485; AAC31807.1; -.

DR PIR; T14271; T14271.

DR HSSP; P16109; 1FSB.

DR MGD; MGI:96083; Nrg1.

DR MGD; MGI:2447063; Odz4.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR001258; NHL.

DR InterPro; IPR006530; YD.

DR Pfam; PF00008; EGF; 3.

DR Pfam; PF01436; NHL; 6.

DR Pfam; PF05593; RHS repeat; 5.

DR SMART; SMO0181; EGF; 6.

DR TIGRFAMS; TIGR01643; YD\_repeat\_2x; 7.

DR PROSITE; PS00022; EGF\_1; 8.

DR PROSITE; PS01186; EGF\_2; 7.

KW EGF-like domain.

SQ SEQUENCE 2825 AA; 313483 MW; A120D98080886032 CRC64;

Query Match 95.8%; Score 14269; DB 11; Length 2825;

Best Local Similarity 93.4%; Pred. No. 0;

Matches 2654; Conservative 47; Mismatches 52; Indels 88; Gaps 4;

QY 1 MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60

Db 1 MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60

QY 61 KDIVPQEAEEFCRTGANFTLRELGLBEVTPPHGTYLRTDGLPQCGYSMGAGSDADMEAD 120

Db 61 KDMVPQEAEEFCRTGNTFTLRELGLGEMTPPHGTYLRTDGLPHCGYSMGAGSDADLEAD 120

QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHEHENTET-----DHPGGL 170

Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHEHENTETGAPLHCSASSTPIEQ 180

QY 165 -----DHPGGL 170

Db 181 SPSPPSPPANESQRRLLGNVQAQTPDSDSEEFVFNFLVKSGSASLGVAANDHPSSL 240

QY 171 QNHARLRTPPPPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDLHSLSGEPAGGAQ 230

Db 241 QNHPRLRTPPPPLPHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDLHSLSGEPAGSAQ 300

QY 231 EPAHAQENWLLNSNIPLETNRLGKQPFGLGTLQDNLIEMDILGASRRHDGAYSDGHFLFKPG 290

Db 301 EPTHAQDNWVLNSKIPVETNRLGKQPFGLGTWQDNLIEMDIFASRRRDGAYSDGHFFFKPG 360

QY 291 GTSPLFCTTSPGYPLTSSVYSPPPRPLPRSTFARPAFNLKKPSKYNKWKCAALSAIVIS 350

Db 361 GTSPLFCTTSPGYPLTSSVYSPPPRPLPRSTFARPAFNLKKPSKYNKWKCAALSAIIS 420

QY 351 ATLVILLAYFVAMHLFGLNWHLQPMEG--QMYEITEDASSWPVPTDVSLYPSGGTGLET 408

Db 421 ATLVILLAYFVAMHLFGLNWHLQPMEGQMOMYEITEDASSWPVPTDVSLYPSGGTGLET 480

QY 409 PDRKKGTTGKPSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRISOVFIDHPVHLKFNV 468

Db 481 PDRKKGAAEGKPSLFPEDSFIDSGEIDVGRRASQKIPPGTFWRISOVFIDHPVHLKFNV 540

QY 469 SLGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSPSHET 528

Db 541 SLGKAALVGIYGRKGLPPSHTQDLDFVELLDGRRLLTQEARSLGTPRQSRGTPVPSPSHET 600

QY 529 GFIOYLDSGIWHLAFYNDGKESEVVSFLTTAIESVDNCPNICYNGDCCISGTCHCFLGFL 588

Db 601 GFIOYLDSGIWHLAFYNDGKESEVVSFLTTAIESVDNCPNICYNGDCCISGTCHCFLGFL 660

QY 589 GPDGGRASCPVLCSGNGQYNMGRCLCHSGWKGAECDFVTNQCIDVACSNHGTCTITGTCIC 648

Db 661 GPDGGRASCPVLCSGNGQYNMGRCLCHSGWKGAECDFVTNQCIDVACSHGTCTIMGTIC 720

QY 649 NPGYKGESCEEVDCMDPTCSGRGVCVRGECHCFVWGVTNCETPRATCLDQCSEHGTFPLP 708

Db 721 NPGYKGESCEEVDCMDPTCSSRGCVCVRGECHCSVWGVTNCETPRATCLDQCSEHGTFPLP 780

QY 709 DTGLCSDPSPWTHDCSIEICAAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHG 768

Db 781 DTGLCNDPSPWTHDCSIEICAAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHG 840

QY 769 TCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHVCVQLGWRGA 828  
Db 841 TCRDGKCECTPGWNGEHCTI-----EGCPGLCNGNGRCTLDLNGWHVCVQLGWRGT 891  
QY 829 GCDTSMETACGSKDNDGDLVDCMDPDCLQLCHINPLCLGSPNPLDIIQETQVPVSQ 888  
Db 892 GCDTSMETGCGDKDNDGDLVDCMDPDCLQLCHVNPLCLGSPDPLDIIQETQAPVSQ 951  
QY 889 QNLHSFYDRIKFLVGRDSTHIIIPGENPDGGHACVIRQVMTSDGTPLVGVNISFVNNPL 948  
Db 952 QNLNSFYDRIKFLVGRDSTHSIPGENPDGGHACVIRQVMTSDGTPLVGVNISFINNPL 1011  
QY 949 FGYTISRQDGSFDLVTNGGSIILRFRERAPFITQEHTLWLPWDRFFVVMETIIMRHEENEI 1008  
Db 1012 FGYTISRQDGSFDLVTNGGSIILRFRERAPFITQEHTLWLPWDRFFVVMETIVMRHEENEI 1071  
QY 1009 PSCDLSNFARPNPVVSPSLTSPASSCAEKGPVPEIQALQBEISISGCKMRLSYLSSRT 1068  
Db 1072 PSRDLSNFARPNPVVSPSLTSPASSCAEKGPVPEIQALQBEIVIAGCKMRLSYLSSRT 1131  
QY 1069 PGYKSVLRISLTHPTIPFNLKMKVHLMVAVEGRFLRKWFAPADLSYFIFWDKTDVYNQKV 1128  
Db 1132 PGYKSVVRISLTHPTIPFNLKMKVHLMVAVEGRFLRKWFAPADLSYFIFWDKTDVYNQKV 1191  
QY 1129 FGLSEAFVSVGYEYESCPDLILWEKRTVLOGYEIDASKLGWSLDKHALNIQSGILHK 1188  
Db 1192 FGLSEAFVSVGYEYESCPDLILWEKRTAVLOGYEIDASKLGWSLDKHALNIQSGILHK 1251  
QY 1189 NGNGENQFVSQOPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDF 1248  
Db 1252 NGNGENQFVSQOPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDF 1311  
QY 1249 NYIRRIFFPSGNVTNILELRNKRDFRHSHPAHKYIYLATDPMSCAVFLSDNSRRRVFKVKT 1308  
Db 1312 NYIRRIFFPSGNVTNILEM-----SHSPAHKYIYLATDPMSCAVFLSDTNSRRRVFKVKT 1364  
QY 1309 VVKDLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTM 1368  
Db 1365 TVVKDLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTM 1424  
QY 1369 IRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNV 1428  
Db 1425 IRRVDQNGIISTLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDNNV 1484  
QY 1429 VLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDE 1488  
Db 1485 VLQISENHQVRIVAGRPMHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDE 1544  
QY 1489 KKNIRIRQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDGYAKDAKLTNPSSLAVCADG 1548  
Db 1545 KKNIRIRQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDGYAKDAKLTNPSSLAVCADG 1604  
QY 1549 ELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDYLY 1608  
Db 1605 ELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFDTSKHLYTQSLPTGDYLY 1664  
QY 1609 NFTYTGDGDIITHIDNNGNMVNVRDSTGMPLWLVPDQGVYVMTGMTNSALKSVTQGH 1668  
Db 1665 NFTYTGDGDIITHIDNNGNMVNVRDSTGMPLWLVPDQGVYVMTGMTNSALRSVTQGH 1724  
QY 1669 ELAMMTYHGNSGLLATKSNENGWTFEYEDSFGRLTNVTFTPTQVSSFRSDTSSVHVQV 1728  
Db 1725 ELAMMTYHGNSGLLATKSNENGWTFEYEDSFGRLTNVTFTPTQVSSFRSDTSSVHVQV 1784  
QY 1729 ETSSKDDVTITTNLSAGAFYITLLQDQVRNSYIIGADGSLRLLLLANGMEVALQTEPHLLA 1788  
Db 1785 ETSSKDDVTITTHLSGSGAFYITLLQDQVRNSYIIGADGSLRLLLLANGMEVALQTEPHLLA 1844  
QY 1789 GTVNPVTKRNVTLPIDNGLNLVWQRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTR 1848  
Db 1845 GTVNPVTKRNVTLPIDNGLNLVWQRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTR 1904

QY 1849 TEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVITYSPGVIAGIQGIMSERMEYD 1908  
Db 1905 TEKIYDDHRKFTLRILYDQAGRPSFWSPSSRLNGVNVITYSPGHIAGIQGIMSERMEYD 1964  
QY 1909 QAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLSSVTMPNVARQTLET 1968  
Db 1965 QAGRITSRIFADGKMWSYTYLEKSMVHLHLSQRQYIFEFDKNDRLSSVTMPNVARQTLET 2024  
QY 1969 IRVSGYRYRNIYQPPPEGNASVIOQDFTEDGHLHHTFYLGTGRVYKYGKLSKLAETLYDTT 2028  
Db 2025 IRVSGYRYRNIYQPPPEGNASVIOQDFTEDGHLHHTFYLGTGRVYKYGKLSKLAETLYDTT 2084  
QY 2029 KVSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIRFRTTEEGMVNARFDYNDNS 2088  
Db 2085 KVSFTYDETAGMLKTIVNLQNEGFTCTIRYRQIGPLIDRQIRFRTTEEGMVNARFDYNDNS 2144  
QY 2089 FRVTSMQAVINETPLPIDLYRYDDVSGKTEQFKFGVIYYDINQIIITTAVMTHTKHEDAY 2148  
Db 2145 FRVTSMQAVINETPLPIDLYRYDDVSGKTEQFKFGVIYYDINQIIITTAVMTHSKHEDAY 2204  
QY 2149 GRMKEVQYEIFRSLMYMTVQYDNMGRVVVKELKVGPYANTTRYSEYDADGQLQTVSIN 2208  
Db 2205 GRMKEVQYETFRSLMYMTVQYDNMGRVVVKELKVGPYANTTRYSEYDADGQLQTVSIN 2264  
QY 2209 DKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDIRDITRLGSDVQYKMDDEGFLRQGGDI 2268  
Db 2265 DKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDLDRITRLGSDVQYKMDDEGSLRQGGDV 2324  
QY 2269 FEYNSAGLLIKAYNRAGSWSVRYRDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYN 2328  
Db 2325 FEYNSAGLLIKAYNRAGSWSVRYRDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYS 2384  
QY 2329 HSSSEITSLYYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYAYGEI 2388  
Db 2385 HSSSEITSLYYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYAYGEI 2444  
QY 2389 YMDTNPNFQIIIGYHGGLYDPLTKLVHMGRDDYDLVLAGRWTSPDHELWKHLSSSNVMPFN 2448  
Db 2445 YMDTNPNFQIIIGYHGGLYDPLTKLVHMGRDDYDLVLAGRWTSPDHELWKRLSSNSIVPFH 2504  
QY 2449 LYMFKNNNPISNSODIKCFMTDVNSWLLTFGFQLHNVI PGYPKPDMDAMEPSYELIHTQM 2508  
Db 2505 LYMFKNNNPISNSODIKCFMTDVNSWLLTFGFQLHNVI PGYPKPDMDAMEPSYELVHTQM 2564  
QY 2509 KTOEWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSQQAPKTKKFASSGSVFGK 2568  
Db 2565 KTOEWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSQQAPETKKFASSGSIFGK 2624  
QY 2569 GVKFALKDGRVTTDIIISVANEDGRRVAAAILNHAHYLENLHFTIDGVDTHTYFVKPGPSEGD 2628  
Db 2625 GVKFALKDGRVTTDIIISVANEDGRRRIAAILNNAHYLENLHFTIDGVDTHTYFVKPGPSEGD 2684  
QY 2629 LAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEEKA 2688  
Db 2685 LAILGLSGRRRTLENGVNVTVSQINTMLSGTRRRYTDIQLQYRALCLNTRYGTTVDEEKV 2744  
QY 2689 RVLELARQRAVRQAWAREQQRLREGEEGLRAWTEGEXQOVLSTGRVQGYDGFVTSVEQY 2748  
Db 2745 RVLELARQRAVRQAWAREQQRLREGEEGLRAWTDGEKQOVLNTRGVQGYDGFVTSVEQY 2804  
QY 2749 PELSDSANNIHFMRQSEMGR 2769  
Db 2805 PELSDSANNIHFMRQSEMGR 2825

RESULT 3  
Q9W7R3  
ID Q9W7R3 PRELIMINARY; PRT; 2824 AA.  
AC Q9W7R3;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Ten-m4.



QY 1547 DGELYVADLGNIRIRFIRKPKPFLNTQNMVELSSPIDQELYLEFDTTGKHLXTQSLPTGDX 1606  
DB 1602 DGELFIADLGNIRIRYVRNKAFLNPLNMYEISSPIDDELYLFDVNASHVFTQSLTTGDX 1661  
QY 1607 LYNFTYTGDGDTLITDNGNMVNRDSTGMPPLWLVPDQGVYVWMTGMTNSALKSVTQ 1666  
DB 1662 LYNFTYSGEGDLSSITDKXKNRVSIRRSTGLPLWLMGPDGQTFWFTGMTNALKSVAQ 1721  
QY 1667 GHELAMTYHNGSGLLATKSNENGWITTFEYDVSFGRLLTNVTFTGQVSSFRSDTDSVHV 1726  
DB 1722 GQEIAMTYHNGSGLLATKSNEDGWSTFEYDNYGRLLTNVTFTGVRSSYRTDSDSTVRV 1781  
QY 1727 QVETSSKDDVTITNLSASGAFYTLQDQVRNSYYIGADGSLRLLLLANGMEVALQTEPHL 1786  
DB 1782 QTEGSKEDITVTNLSASGTFYTLMDQVKNYSYYIGLDSURLVLVLANGMEVSLHTEPHL 1841  
QY 1787 LAGTVNPTVGKRNVTLPIDNGLNLVWQRKEQARGQVTFGRRLRVHNRNLLSLDFRV 1846  
DB 1842 LSGTVNPTISKRNVTLPIDNGLNLVWQRKEQARGQVTVGRRLRVHNRNLLSMDFRV 1901  
QY 1847 TRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTSYSGYIAGIQRGIMSERME 1906  
DB 1902 TRTEKVYDDHRKFTLRILYDHAGRPTLWAPSSRLNGVNVTSYSGYIAGIQRGIMSERME 1961  
QY 1907 YDQAGRITSRIPADGKTWSYTYLEKSMVLLLSQRQYIFEFDKNDRLSSVTMPNVARQTL 1966  
DB 1962 YDQNGRITSKIPADGKWSYTYLEKSMVLLLSQRQYIFEFDKNDRLSSVTMPNVARQTL 2021  
QY 1967 ETIRSVGYRNYIQQPEGNASVIQDFTEDGHLHTFYLGTGRVYKYGKLSKLAETLYD 2026  
DB 2022 ETTRSIGYRNYRYPPEGNATVLQDYSDEGLLQTIHQGTGRVYKYGKLSRLLLELYD 2081  
QY 2027 TTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEEGMVNARFDYND 2086  
DB 2082 TTRIAFSYDESAGMLKTIVGLQSEGFACTIRYQIGPLIDRQIFRFTTEEGMVNARFDYND 2141  
QY 2087 NSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHF 2146  
DB 2142 NSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHF 2201  
QY 2147 AYGRMKEVQYEIERSLMYMTVQYDNMGRVVKELKVGPYANTRYSYEYDADGQLQTVS 2206  
DB 2202 AYGRVKEVQYEIERSLMYMTVQYDNMGRVVKELKVGPYANTRYAYEYDADGQLQTVS 2261  
QY 2207 INDKPLWRYSDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDDVQYKMDDEGFLRQGG 2266  
DB 2262 INDKPLWRYSDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDDVQYRLDEDEGFLRQGN 2321  
QY 2267 DIFEYNSAGLLIKAYNRAGWSVRYDYDGLGRRVSSKSSHHLQFFYADLTNPTKVTHL 2326  
DB 2322 DFFEYNSAGLLVKTYNKVGWTKYKYDGLGRRVSSRSTQGHHLQFFYADLSSPTRVTHM 2381  
QY 2327 YNHSSEITSLYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTA 2386  
DB 2382 YNHSSEITSLYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGAGLMIKQILHTAF 2441  
QY 2387 EIYMDTNPFIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHMLWKHLSSSNVMP 2446  
DB 2442 EYLDNSNPSFQVIGYQGGLYEPLTKLVHMGRRDYDVLAGRWTSPDHDIRKRLNSDNIVP 2501  
QY 2447 FNLVFMKNNPLNSQDIKCFMTDVNSWLLTFGQLHNVIPGPKPDMAMDAMEPSYELIHT 2506  
DB 2502 FNLVFMKNNPLNSQETKCYMTDVNSWLVTFGQLYNVIPGVRKPVTDAMEPSYELVHT 2561  
QY 2507 QMKTQEWDSKILGVQCEVQQLKAFVTLERFDQLYGSTITSCQAPKTKKFASSGSVF 2566  
DB 2562 QIKTQEWDSKSVLGVQCEVQQLKSFVRLERFGQIYASDSGCPPTPLHTLFATGTSLF 2621  
QY 2567 KGKVKFALKDGRVTTDIIISVANEDGRRAAAILNHAHYLENLHFTIDGVDTHYFVKPGPSE 2626  
DB 2622 KGKVKVAIREGRVADIIISLANEDGRRIAALDKASYLQDLHFTIAGLDTHYFVKSGLVE 2681  
QY 2627 GDLAIGLGGRRTELVNGVNVTSQINTVLNGRTRRYTDIQLQYCALCLNTRYGTTLDEE 2686

DB 2682 GDSLGLGMTVQRTLETGVNVTVSQVNMVGLGRRSRRTIDIQMQYGTLSLNVRYGSSVDEE 2741  
QY 2687 KARVLELARQRAVRQAWAREQQRLRECEEGELRAWTEGKQQLVLTGRVQGYDGFVTSVE 2746  
DB 2742 KVRVLELARQRAVATAWAHERHRLRQEGEGSRATDGERQQLSSGRVQGYEGFYIVSVD 2801  
QY 2747 QYPELSDSANNIHFMRQSEMGR 2769  
DB 2802 QFPELTDNINNHFWRQTEMGR 2824  
PRT; 2715 AA.  
Q9WTS6 PRELIMINARY;  
Q9WTS6;  
01-NOV-1999 (TREMELrel. 12, Created)  
01-NOV-1999 (TREMELrel. 12, Last sequence update)  
01-OCT-2003 (TREMELrel. 25, Last annotation update)  
Ten-m3.  
OD23 OR TEN-M3.  
Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=Balb/c; TISSUE=Brain;  
RA Cohashi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,  
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;  
RT "Mouse Ten-m/Odz is a new family of dimeric type II transmembrane  
proteins expressed in many tissues."  
RL J. Cell Biol. 0:0-0(1999).  
DR EMBL; AB025412; BAA77398.1; -.  
DR HSSP; P01492; 1VNB.  
DR MGD; MGI:1345183; Odz3.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004623; F:phospholipase A2 activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0016042; P:lipid catabolism; IEA.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001258; NHL.  
DR InterPro; IPR001211; PhospholipaseA2.  
DR InterPro; IPR006530; YD.  
DR Pfam; PF00008; EGF; 3.  
DR Pfam; PF01436; NHL; 6.  
DR Pfam; PF05593; RHS repeat; 6.  
DR PRINTS; PR00011; EGFLAMININ.  
DR SMART; SM00181; EGF; 6.  
DR TIGRFAMs; TIGR01643; YD\_repeat\_2x; 6.  
DR PROSITE; PS00022; EGF\_1; 8.  
DR PROSITE; PS01186; EGF\_2; 7.  
KW EGF-like domain.  
SQ SEQUENCE 2715 AA; 303062 MW; 598F46A77334C2E1 CRC64;  
Query Match 68.7%; Score 10227; DB 11; Length 2715;  
Best local Similarity 66.7%; Pred. No. 0;  
Matches 1864; Conservative 370; Mismatches 457; Indels 102; Gaps 19;  
QY 1 MDVKERKPYPSLTR-RRDAERRYTSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57  
DB 1 MDVKERRPYCSLTKSRRKERRYTNSSADNEECRVPTQKSYSSSETLKAFDHDYSRLLYG 60  
QY 58 SRVKDIVPQEAEEFCRTGANFTLRGLLEEVTPPHGTLTYRTDIGLPQCQGYSMGAGSDADM 117  
DB 61 NRKVDLVHREADEYTRQGNFTLRQLGVCESTRRGVAFCAEMGLPHRGYSISAGSDAT 120  
QY 118 EADTVLSPHPVRLWGRSTRSGRSCSLSSRANSNLTLTDTHEH---TETHPGGLQNH 174  
DB 121 ENEAVMSPEHAMRLWGRGVKSGRSCSLSSRSNSALTTLTDTHEHNRSDSESEQPSNPGQP 180

Qy	175	RLRTPPPPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDSLSGEPAGAQ----	230
Db	181	TLQLPP--SHKQHPAQHH-PSITSLNRNSLTNRNRNQSPAP-----PAALPAELQTT	229
Qy	231	-EPAHAQENWLLNSNIPILETRNLGKQPFGLGTLDNLIEMDILGASRRHDGAYSDDGHFLFKP	289
Db	230	PESVQLQDSWVLGSNVPLESR-----HFLFKT	256
Qy	290	G-GTSPLFCTTSPGYPLTSTSVYSPPPRPLPRSTFARPAFNLKKPSKYCNWKAALSAIV	348
Db	257	GTGTTPLFSTATPGYTMASGSVYSPTRPLPRNTLSRSFAFKKSSKYCSWRCTALCAVG	316
Qy	349	ISATLVILLAYFVAMHLFGLNWHLQPMEGOMYE--ITEDTASSWPVPTDVSLYPSGGTG	405
Db	317	VSULLAILLSYFIAMHLFGLNWHLQQTENDTFENGKVNSDT---VPTNTVSLPSG---	368
Qy	406	LETPDRKGKGTGKPPSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHPVHLK	465
Db	369	-----DNCKLGGFTHENNTIDSGELDIGRRAIQEVPPGIFWRSQLFIDQPQFLK	417
Qy	466	FNVSGLKAALVGIYGRKGLPPSHTQDFVVELLDGRRLLTQEARSLGEGTTPRQSRGTVPSS	525
Db	418	FNISLQKDALIGVYGRKGLPPSHTQYDFVELLDGSRLLIAREQRLNLSERAGRQARSVL	477
Qy	526	HETGFIQYLDSDGIWHLAFYNDGKSESEVVSFLTIAESVDNCPNSCYNGNGDCISGTCHCFL	585
Db	478	HEAGFIQYLDSDGIWHLAFYNDGKNPEQVSNFTIVIESVVECPRNCHNGGECVSGTCHCFF	537
Qy	586	GFLGPDCCRACPVLCSGNGOYMKGRCLCHSGWGAECDDVPTNQCIDVACSNHGTCTGT	645
Db	538	GFLGPDCCRACPVLCSGNGOYSGKGRCLCFSGWKGTEDCVPTTQCIDPQCGGRGICIMS	597
Qy	646	CICNPGYKGESCEEVDCMDPTCSGRGVVGRGECHCFVGWGTNCETPRATCLDQCSGHGT	705
Db	598	CACNSGYKGENCEEADCLDPGCSNHGVCIHGECHCNPGWGGSNCEILKTMCADQCSGHGT	657
Qy	706	FLPDTGLCSDEPSWTGHDCSIICAADCGHGVCVGGTCRCEGDGWMGAACDQACHPRCA	765
Db	658	YLQESGCTCDPNWTGPDCSNEICSVDCGSHGVCMGGSCRCCEGWTGPACNQACHPRCA	717
Qy	766	EHGTCRDGKCEGPGWNGEHCTIAHYLDRVVEGCPGLCNGNGRCTLDLNGWHCVCQLGW	825
Db	718	EHGTCXDKGCEGSGWNGEHCTIAHYLDKIVKEGCPGLCNSNGRCTLDQNGWHCVCQPGW	777
Qy	826	RGAGCDTSMETACGDSKNDGDGLVDCMDPDCCLQPLCHINPLCLGSPNPLDIQETQVP	885
Db	778	RGAGCDVAMETLCTDSKDNEGDLIDCMDPDCCLQSSCQNPYCRGLPDPQDIISQSLQT	837
Qy	886	VSQNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRQVMTSDGTPLVGNISFVN	945
Db	838	PSQQAAKSFYDRISFLIGSDSTHVLPGESPFNKSLASVIRQVLTADGTPLIGVNVSFLLH	897
Qy	946	NPLFGYTTISRDGSFDLVNTGGISIIILRFERAPFITQHTLWLPWDRFFVMTIIMRHEE	1005
Db	898	YSEYGYTTITRODMFDLVANGGASLTIVFERSPLTQYHTVWIPWNVFVYMDLVMKKEE	957
Qy	1006	NEIPSCDLSNFARPNPVVSPLTSFASSCAEKGPVPEIQALQEEISISGCKMRLSYLS	1065
Db	958	NDIPSCDLSGFVRPSPPIIVSSPLSTFFRSSPEDSPIPETQVLHEETTPGTDLKLSYLS	1017
Qy	1066	SRTPGYKSVLRI SLTHPTIPFNLMKVHLMVAEGELFRKWFAAAPDLSYFIWKTIDVYN	1125
Db	1018	SRAAGYKSVLKITMTQAVIPFNLMKVHLMVAUVGELFQKWFPASPNLAYTFIWDKTDAYN	1077
Qy	1126	QKVFGLSAEFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHHALNIQSGI	1185
Db	1078	QKVYGLSEAVVSVGYEYESCLDLTLWEKRTAVLQGYELDASNMGWTLTDKHHVLDVQNGI	1137
Qy	1186	LHKGNGENQFVSQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCSDGSLYV	1245
Db	1138	LYKGNGENQFISQPPVVSIMGNRRRSISCPSCNGQADGNKLLAPVALACGIDGSLYV	1197
Qy	1246	GDFNYIRRIFFPSGNVTNILELRNKDFRHSHPAHKYVYLATDPMGSAVFLSDSNRRRVFKI	1305

Db	1198	GDFNYVRIFPSGNVTSVLELRNKDFRHSNPAHRYYLATDPVTGDLVYSDTNRRIYRP	1257
QY	1306	KSTVVVKDLVKNSEVVAGTGDOCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVD	1365
Db	1258	KSLTGAKDLTKNAEVVAGTGEQCLPFDEARCGDGGKAVEATLMSPKGMAIDKNGLIYFVD	1317
QY	1366	GTMIIRIDONGIISTLLGSNDLTSARPLSCDSVMDISOVRLEWPTDLAINPMDNSLYVLD	1425
Db	1318	GTMIKVDQNGIISTLLGSNDLTSARPLTCDTSMHISOVRLEWPTDLAINPMDNSIYVLD	1377
QY	1426	NNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAE	1485
Db	1378	NNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAVSYSGLYITE	1436
QY	1486	TDEKKINRIQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLNTPSSLAVC	1545
Db	1437	TDEKKINRIQVTTDGEISLVAGIPSECDCKNDANCDCYQSGDGYAKDAKLNAPSLLAAS	1496
QY	1546	ADGELYVADLGNIRIRFIRKKNPFLNTQNMVELSSPIDQELYLFDTTCCKHLYTQSLPTGD	1605
Db	1497	PDGTLYIADLGNIRIRAVSKNPLLNMMNFYEVASPTDQELYIFDINGTHQYTVSVLVTGD	1556
QY	1606	LYXNFTYTGDGDIITLITDNNGNMNVRRDSTGMPLWLVPDQVYVWVTGMTNSALKSVTT	1665
Db	1557	LYNFSYNDNDVTAVTDNGNTLRIRRDPNRMPVRVVPDNQVWLVTIGTNGCLKSMTA	1616
QY	1666	QGHELAMMTYHGNSGLLATKSNENGWTTFFEYDSFGRLTNVTFPTGVSSFRSDTDSVVH	1725
Db	1617	QGLELVLFYHGNSGLLATKSDETGWTTFDDYDSEGRLTNVTFPTGVVTLNHGDMDKAIT	1676
QY	1726	VQVETSSK-DDVTITTNLSAGAFYTLLODQVRNSYYIGADGSLRLLLLANGMEVALQTEP	1784
Db	1677	VDIESSREEDVSITNLSSIDSFYTMVQDQLRNSYQIGYDGLSRIFVASGLDSHYQTEP	1736
QY	1785	HLLAGTVNPVTKRNVTLPIDINGLNLVEWRQKEQARGQVTFGRRRLVRVHNRNLLSLDFD	1844
Db	1737	HVLAGTANPIVAKRNMVTLPGENGQNLVEWRFRKEQAQGVKVNFGRLVRNGRNLLSVDFD	1796
QY	1845	RVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTYSPGGYIAGIQRGIMSER	1904
Db	1797	RTTKTEKIYDDHRKFLLRIAYDTSGHPTLWLPSSKLMANVTYSSGTGQIASIQRTTSEK	1856
QY	1905	MEYDQAGRITSRIFADCKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDLSSVTMPNVARQ	1964
Db	1857	VDYDSQGRIVSRVFADGKTWSYTYLEKSMVLLHLSQRQYIFEYDMWDLBSAITMPSVARH	1916
QY	1965	TLETIRSVGYRNIYQPPEGNASVIOFDETDGHLHLHTFYLGTGRRVYKYGKLSKLAETL	2024
Db	1917	TMQTIRSIGYRNIYNPPESNASIITDYNBEGLLLTQTAFLGTSRRVLFKYRRQTRLSEIL	1976
QY	2025	YDTTKVSFTYDETAGMLKKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTBEGMVNARFDYN	2084
Db	1977	YDSTRVSFTYDETAGVLKTVNLQSDGFTCTIRYRQIGPLIDRQIFRFSBEGMVNARFDYS	2036
QY	2085	YDNSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKH	2144
Db	2037	YDNSFRVTSMQAVINETPLPIDLYQFDDISGKVEQFGKFGVIYYDINQIISTAVMTYTKH	2096
QY	2145	FDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVVKELKVGPYANTTRYSEYDADGQLQT	2204
Db	2097	FDAHGRIKEIQYEIFRSLMYWITIYDNMGRVTKREIKIGPFANTTKYAYEYDVGQLQT	2156
QY	2205	VSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFLRQR	2264
Db	2157	VYLNEKIMWRYNDLNGNLHLLNPPSSSARLTPLRYDLDRDIRLGDVQYRLDEDEGFLRQR	2216
QY	2265	GGDIFEYNSAGLLIKAYNRAGSWSVRYRDGLGRRVSSKSSHHLQFFYADLTNPTKVT	2324
Db	2217	GTEIFEYSSKGILLTRVYSGSGWTVIYRDGLGRRVSSKTSLGQHLQFFYADLTYPTRIT	2276
QY	2325	HLXNHSSSEITSLYYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTA	2384

Db 2277 HVYNHSSSEITSYYDLQGLHFAFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLKQTQYTA 2336  
QY 2385 YGEIYMDTNPFIIGYHGLYDPLTKLVHMGRRDYDVLGRWTS PDHELKWLSSNV 2444  
Db 2337 YGEIYFDSNVDFQVIGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDIEIWKRI-GKDP 2395  
QY 2445 MPENLYMFKNPNISNSQDIKCFMTDVNSWLLTFGQLHNVI PGYPKPDMDAMEPSYELI 2504  
Db 2396 APENLYMFRNNPASKIHVDKDYITDVNSWLVTFGFHLHNAIFGPPVPKFDLTPEPSYELV 2455  
QY 2505 HTQMKTQEWDSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSCQAPKTKK----- 2558  
Db 2456 ----KSQQWEDVPPIFGVQQQVARQAKAFSLGKMAEV-----QVSRKAGAEQSW 2502  
QY 2559 --FASSGSVFGKGVKALKDGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDT 2616  
Db 2503 LWFATVKSILIGVMLAVSQGRVQTNLNLANEDCIKVAAVLNNAFYLENLHFTIEGKDT 2562  
QY 2617 HYFVKPGPSEGLAILGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDIOLQYGALCLN 2676  
Db 2563 HYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVNGRTRRFADVEMQFGALALH 2622  
QY 2677 TRYGTTLDEEXARVLELARQAVRQAWAREQQRLREGEEGLRAWTEGEKQQVLSTGRVQG 2736  
Db 2623 VRYGMTLDEEXARILEQARQALARAWAREQQRVRDGEGARLWTEGEKRLLSAGKVQG 2682  
QY 2737 YDGFVVISVEQYPELSDSANNIHFMRQSEMGR 2769  
Db 2683 YDGYVLSVEQYPELADSANNIQLRQSEIGKR 2715

RESULT 5

Q804R1 PRELIMINARY; PRT; 2192 AA.  
ID Q804R1  
AC Q804R1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE SI:Cl97B11.1 (odd Oz/ten-m homolog 4) (Fragment).  
GN OD24.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Babbage A.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL807754; CAD61260.1; -.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001258; NHL.  
DR InterPro; IPR006530; YD.  
DR Pfam; PF01436; NHL; 4.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR SMART; SM00181; EGF; 8.  
DR TIGRFAMs; TIGR01643; YD repeat\_2x; 5.  
DR PROSITE; PS00022; EGF\_1; 8.  
DR PROSITE; PS01186; EGF\_2; 7.  
FT NON TER 1  
SQ SEQUENCE 2192 AA; 243920 MW; 063A4B36DA3718A5 CRC64;

Query Match 67.0%; Score 9968; DB 13; Length 2192;  
Best Local Similarity 80.3%; Pred. No. 0;  
Matches 1774; Conservative 242; Mismatches 176; Indels 16; Gaps 2;  
QY 562 SVDNCPSCNCGNGDCISGTCFLGFLGPDCCGRASCPVLCSGNGQYMKGRCLCHSGWKA 621  
Db 1 SIDDCPSNCFNGDCVSGNCHCFPGFRGPDCSRASCPVLCSGNGQYLGKRCMCHSGWKGS 60

QY 622 ECDVPTNQCIDVACSNHGTCTITGTCTICNPGYKGESCEEVDCMDPTCSGRGVCVRGECHCF 681  
Db 61 ECDVPTNQCIDITCSGHGTCIVGTCTICNPGYKGENCEEVDCLDPTCSGRGVCVRGECHCF 120  
QY 682 VGGGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPSWTGHDCSIEICAADCGHGVCVG 741  
Db 121 VGGGPGCESPRASCMQCSGHGSLADTNTCNDHNWTHGDCSTELCAADCGHGICVA 180  
QY 742 GTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSPGNGEHCTIAHYLDRVVKEGCP 801  
Db 181 GSCRCDEGWMGTGCEQACHPRCSEHGTCKDGKCECSPGNGEHCTI-----EGCP 231  
QY 802 GLCNGNGRCTLDLNGWHVCQLGWEGAGCDTSMETACGDSKNDGDGLVDCMDPDCCLOP 861  
Db 232 GLCNGNGRCTLGNNGWYCVQQLGWEGAGCDTSMETACSDGKNDGDGLTDCMDPDCCLOA 291  
QY 862 LCHINPLCLSPNPLDIIQETQVPSQQNLHSFYDRIKELVGRDSTHIIIPGENPFDGHA 921  
Db 292 SCHTTSLCVSPDPLDIIQETQISSLSLTQSFYQRIHFLVGRDSTHIVPGVNPFDGIHA 351  
QY 922 CVIRGOVMTSDGTPLVGVNISFVNPLFGYTI SRQDGSFDLVNNGGSIILRFERAPFIT 981  
Db 352 CVIRGOVMTSDGTPLVGVNISFINKPAYGTYITRQDGSFDLVNNGGVAIGLRFERAPFIT 411  
QY 982 QEHTLWLPWDRFFVPMETIIMRHENEIIPSCDLNFAFNPVPSPLTSFASSCAEKGP 1041  
Db 412 QEHTLWLPWGRFFVMDTIIVMRHEVNDIPSCDLSSFTTRPMPVLPAPLTAFAGTCPERGIV 471  
QY 1042 VPEIQALQEEISISGCKMRLSYLSRTPPGYKSVLRISLTHPTIPFNLKMLHMAVEGRL 1101  
Db 472 VPEIQTLQEEVRIPGTDMLRGLYSRTSGYKSLRLITLTHSTIPFSLMKVHLMVAVEGRL 531  
QY 1102 FRKWFAAAPDLSYFIWDKTDVYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTTVLQGY 1161  
Db 532 FRKWFAAPNLSYDFVWDKTDVYSQKYGLSEAFVSVGYEYESCPDLILWEKRTAVLQGY 591  
QY 1162 EIDASKLGGWLDKHALNIQSGLHKGNGENQFVSQQPPVIGSIMGNRRRSISCPSCN 1221  
Db 592 ETTASNLGGWSVDKHALNIQSGLHKGNGENIFISQQPPVIGSIMGNRRRSISCPSCN 651  
QY 1222 GLADGNKLLAPVALTCGSDGLYVGFNYIRRIFFPSGNVTNILELRNKDFRSHSHPAHY 1281  
Db 652 GLADGNKLLAPVALACGSDGLYVGFNYVRRIFTTGNVTSLVEL-----SNSPAHY 704  
QY 1282 YLATDPMGSAVFLSDSNRRRVEKTKSTVWVKLVKNSEVAVAGTDQCLPFDDTRCGDGGK 1341  
Db 705 YLATSPVSGWLYLSDTSSRKVKVKSILYAVKDVAKNLELVACTGQCLPYDETRCGDGGK 764  
QY 1342 ATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIISTILGSNDLTSARPLSCDSYMDI 1401  
Db 765 AVEATLTNPRGITVDKYGVFFVDGTMIRRIDQNGIISTILGFNDLTSARPLSCDSYMDI 824  
QY 1402 SQVRLEWPTDLAINPMDNSLYLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKV 1461  
Db 825 SQVRLEWPTDLAVSPMDNSLYLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLVSKI 884  
QY 1462 AIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTSGEISLVAGAPSGCDCXNDANC 1521  
Db 885 AIHATLESANALAVSHNGLLYIAESDEKKINRVQVSTNGEISLLAGAPSGCDCXNDANC 944  
QY 1522 DCFSGDDGYAKDAKLTNPSSLAVCADGELYVADLGNIRIRFIRKKNKPFNLTONMYELSSP 1581  
Db 945 DCYSGDDGYAKDAKLNAPSSLAVSPDGLFIADLGNIRIRYVRRNKAFNLPLNMYEISSP 1004  
QY 1582 IDQELYLEDTTGKHLTYQSLPTGDIYLYNFTYTGDGDIITLTDNNGNMVNRDRDSTGMLPW 1641  
Db 1005 IDDELYLFDVNASHVFTQSLTTGDIYLYNFTYSGEGDLSITDKNKNRVSIRRDSTGLPLW 1064  
QY 1642 LVVPDQGVYVWMTGNTSALKSVTTQGHLELAMMTYHNSGLLATKSNENGTWTFYEDSFG 1701  
Db 1065 LMGPDGGQTFWFTMGTTNNALKSVAAQGGQIEIAVMTYHGSSGLLATKSNEDGWSTFYEDNYG 1124







QY 1205 SIMGNRRRSISPCSCNGLADGNKLLAPVALTCGSDGLYGVDFNIRRIFFSPGNVTNII 1264  
Db 1241 SIMGNRRRSISPCSCNGLAEGNKLAPVALAVGIDGSLFVGDFNIRRIFFSPRNVTSL 1300  
QY 1265 ELRNKDFRHSHPAHKYVYLATDPMSCGAVFLSDSNSRRVFKIKSTVVVKDLVKNSEVAVGT 1324  
Db 1301 ELRNKEFKHSNNPAHYVYLAVDPVSGSLYVSDTNSRRIYKVKSLTGKDLAGNSEVAVGT 1360  
QY 1325 GDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIISTLLGS 1384  
Db 1361 GEQCLPFDEARCGDGGKAVDATLMSPRGLAVDKYGLMYFVDATMRKVQDQNGIISTLLGS 1420  
QY 1385 NDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRIVAGR 1444  
Db 1421 NDLTAVRPLSCDSMDVSQVRLEWPTDLAVDPMDNSLYVLENNVILRITENHQVSIAGR 1480  
QY 1445 PMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLVYIAETDEKKINRIRQVTTSGEIS 1504  
Db 1481 PMHCQVPGID-YSLSKLAIHSALESASAIASHTGVLVYISETDEKKINRLRQVTTNGEIC 1539  
QY 1505 LVAGAPSGCDCKNDANCDGFCGDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRIRFIR 1564  
Db 1540 LLAGAASDCDCRNDVNCNCYSGDDGYATDAILNSPSSLAVAPDGTIYIADLGNIRIRAVS 1599  
QY 1565 KNKPFNTQNMVELSSPIDQELYLFDTTGKHLYTQSLPTGDYLYNFTYTGDGITLITDN 1624  
Db 1600 KNRPILNSFNQYEAASPGEQELYVFNADGTHQYTLTSLVTGEYLYNFTYSSDNDVTEVMD 1659  
QY 1625 NGNMVNRRDSTGMPLWLVPDGVVYVWVTMGNTSALKSVTTQGHELAMMTYHGNISGLLAT 1684  
Db 1660 NGNSLKVRDDASGMPRHLLMPDNQIVTLAVGTNGGLKLVSTQTLGLMTYNGNSGLLAT 1719  
QY 1685 KSNENGWTTTFEYDYSFGRLTNTVTPFTGQVSSFRSDTSSVHVQVETSSK-DDVTITNLS 1743  
Db 1720 KSDGTGTTTFDYDHEGRLTNVTRPTGVTVTSLSHREMEKSITIDIENSNRDDDVITNLS 1779  
QY 1744 ASGAFYTLQDQVNSYIIGADGSLRLILLANGMEVALQTEPHLLAGTVNPTVGKRVNVLTP 1803  
Db 1780 SVEASYTVVQDQVNSYQLCNGNGLTRVMYANGMSISFHSPEHVLAVGTPTIGRCNISLP 1839  
QY 1804 IDNGLNLVEMRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRI 1863  
Db 1840 MENGLNSIEWLRKEQIKGKVTVFGRKLVRHGRNLLSIDYDRNIRTEKIYDDHRKFTLRI 1899  
QY 1864 LYDQAGRPSLWSPSSRLNGVNVVTVSPGGYIAGIQRGIMSERMEYDQAGRITSRIFADGKT 1923  
Db 1900 IYDQLGRPFLWLPSSGLAANVSVYFFNGELAGLQRGAMSERITIDKQGRIRSRMFADGKV 1959  
QY 1924 WSYTYLEKSMVLLHLSORQYIFEFKNDRLSSVTMPNVARQLETIRSVGYRNIYQPPPE 1983  
Db 1960 WSYTYLEKSMVLLLQSORQYIFEYDSSDRHLHAVTMPSVARHSMSTHTSVGIYRNIYNPPE 2019  
QY 1984 GNASVIQDFTEDCHLLHTFYLGTGRRVYIKYGLSKSLAETLYDTTKVSFTYDETAGMLKT 2043  
Db 2020 SNASVIFDYSDDGRILKTSFLGTGRQVYKYGLSKSLSEIVYDSTAVTFGYDETTGVLM 2079  
QY 2044 INLQNEGFTCTIRYRQIGPLIDRQIFRPTTEEGMVNARFDYNY-DNSFRVTSMQAVINETP 2102  
Db 2080 VNLQSGGFSCTIRYRKIGPLVLDKQIYRFPSEEGMVNARFDYTYHDSNFRIASIKPIISETP 2139  
QY 2103 LPIDLRYDDVSGKTEQFGKFGVYYVDINQIITAVMTHTKHFDAYGRMKEVQYEIFRSL 2162  
Db 2140 LPVDLYRYDEISGKVEHFGKFGVYYVDINQIITAVMTLSKHFDTHGRIVEQYEMFRSL 2199  
QY 2163 MYWMTVQYDNMGRVVKELKVGYPYANTTRYSEYDADGQLQTVSINDKPLWRYSYDLNGN 2222  
Db 2200 MYWMTVQYDSMGRVTKRELKLGYPYANTTRYKYTYDYGDLQSLQSAVANDRPTWRYSYDLNGN 2259  
QY 2223 LHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDGDFLQRGGDIFEYNSAGLLIKAYN 2282  
Db 2260 LHLNPGNSVRLMPLRYDLRDRITRLGDIPIKIDDDGFLCQRGSDVFEYNSKGLLTRYN 2319  
QY 2283 RAGSWSVRYRYDGLGRRVSSKSSHHLQFFFYADLTNPTKVTHLYNHSSEITSLYYDLQ 2342

Db 2320 KANGWNVQYRYDGLGRRASCKTNLGHHLQYFYADLHNPTRVTHVYNSNSEITSLYYDLQ 2379  
QY 2343 GHLFAMELSSGDEFYIACDNIGTPLAYFSCTGLMIKQILYTAIGEIMYMDTNPFIQIIIGY 2402  
Db 2380 GHLFAMESSGEEYVVASDNTGTPLAVFSINGLMIKQLQYTAIGEIYYDSNPDFQLVIGF 2439  
QY 2403 HGGLYDPLTKLVHMRDRDYDLAGRWTSPDHKLWHLSSSNVMPFNLYMFKNNPISNQ 2462  
Db 2440 HGGLYDPLTKLVHFTQRDYLGRWTSPDYTMWKNIGREPA-PFNLYMFKSNPLSNEL 2498  
QY 2463 DIKCFMTDVNSWLLTFGQLHNVIPGYPKPMDAMEPSYELIHTQMKTQEWDNKSILGV 2522  
Db 2499 DLKNYVTDVKSMLVMFGQLSNIIPGFPRAKMYFVSPPYEL----TESQACENGQLITGV 2554  
QY 2523 QCEVQKQLKAFVTLERFDQLYGSTITSQQQAPKTKK----FASGGSVFGKGVKALKDGR 2578  
Db 2555 QOTTERHNQAFMALE-----GVISKRLHASIREKAGHWFATSTPIIGKIMFAVKGR 2608  
QY 2579 VTTDIISVANEDGRRVAAAILNHAHYLENLHFTIDGVDTHYFVKPSPSEGLAILGLSGGR 2638  
Db 2609 VTTGISSIAATDSDRKIASVLNSAHYLEKMHYSIEGKDTHYFVKIGSADSDLVTLAMTSGR 2668  
QY 2639 RTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYG---TTLDEEKARVLELAR 2695  
Db 2669 KVLDSGVNVTVSQPTLLINGRTRRTNIEFYQYSTLLINIRYGLTADTLDEEKARVLDQAR 2728  
QY 2696 QRAVRQAWAREQORLREGEGLRAWTEGEKQOVLSTGRVQGVDFVVISVEQYPELSDSA 2755  
Db 2729 QRALGSAWAKEQQKARDGREGSRVWTDGEXQQLLNTGRVQGVGYGYVLPVEQYPELADSS 2788  
QY 2756 NNHFMRQSEMGRR 2769  
Db 2789 SNIQFLRQENMGKR 2802

RESULT 8

Q9WTS5 PRELIMINARY; PRT; 2764 AA.  
AC Q9WTS5;  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Ten-m2.  
GN ODZ2 OR TEN-M2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Balb/c; TISSUE=Brain;  
RA Ohashi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,  
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;  
RT "Mouse Ten-m/Odz is a new family of dimeric type II transmembrane  
RT proteins expressed in many tissues."  
RL J. Cell Biol. 0:0-0(1999).  
DR EMBL; AB025411; BAA77397.1; -.  
DR HSSP; P35555; IEMN.  
DR MGD; MGI:1345184; Odz2.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001258; NHL.  
DR InterPro; IPR006530; YD.  
DR Pfam; PF00008; EGF; 4.  
DR Pfam; PF01436; NHL; 6.  
DR Pfam; PF05593; RHS repeat; 6.  
DR PRINTS; PR00011; EGFLAMININ.  
DR SMART; SM00181; EGF; 5.  
DR TIGRFAMs; TIGR01643; YD\_repeat\_2x; 7.  
DR PROSITE; PS00022; EGF\_1; 8.



Db 1959 VARHSMSTHTSIGYIRNIYNPPESNASVIFDYSDDDRILKTSFLGTGRQVFKYKGLSKL 2018

QY 2021 AETLYDTTKVSTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNAR 2080

Db 2019 SEIYDSTAFTGYDETTGVLMVNLQSGGFSCTIRYRKVGLVDKQIYRFSEEGMINAR 2078

QY 2081 FDYNY-DNSFRVTSMAVINETPLPIDLYDDVSGKTEQFGKFGVIYDINOIITAVM 2139

Db 2079 FDYTYHDNSPRIASIKPVISETPLPDLYRYDEISGKVEHFGKFGVIYDINOIITAVM 2138

QY 2140 THTKHFDAYGRMKEVQYEIFRSLMYMTVQYDNGMGRVVKELKVGYPYANTTRYSVEYDAD 2199

Db 2139 TLSKHFDTHGRIKEVQYEMFRSLMYMTVQYDSMGRVVKELKVGYPYANTTKYTYDYDGD 2198

QY 2200 GOLQTVSINDKPLWRYSDYDLNGLHLLSPGNSARLTPLRYDIRDITRLGDVQVKMDEG 2259

Db 2199 GOLQSVAVNDRPTWRYSDYDLNGLHLLNPGNSARLMLPLRYDLDRITRLGDVQVKIDDDG 2258

QY 2260 FLRQGGDIPEYNSAGLLIKAYNRAGSWSVRYRYDGLGRVSSKSSHSHLQFFYADLTN 2319

Db 2259 YLCQSGSDIFEYNSKGLLTRAYNKASGWSVQYRYDGVGRASYKTNLGHLLQYFYSDLHN 2318

QY 2320 PTKVTHLYNHSSEITSLYLDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQ 2379

Db 2319 PTRITHVYHNSSEITSLYLDLQHLFAMESSSGEEYYVASDNTGTPLAVYSINGLMIKQ 2378

QY 2380 ILYTAYGEIYMDTNPFIQIIIGHGGLYDPLTKLVHMGRRDYDLVLAGRTWSPDHHLWKHL 2439

Db 2379 LQYTAYGEIYDSDNPDFQMVIGFHGGLYDPLTKLVHFTQRDYDVLAGRWTSPDYTWARNV 2438

QY 2440 SSSNVMPENLYMFKNNPISNQDIKCFMTDVSLLTFGQLHNVIPGYPKPDMNAMEP 2499

Db 2439 GKEPA-PENLYMFKNNPLSNELDLKNYVTDVXSWLVMFGQLSNIIPGFPRAKNYFVPP 2497

QY 2500 SYELIHTQMKTQEWDNKSIILGVQCEVQKQLKAFVTLERFDQLYGSTITSCQAPKTKK- 2558

Db 2498 PYELSESQAS----ENGQLITGVQOTTERHNOAFLALE-----GQVITKKLHASIREKA 2547

QY 2559 ---FASGSGVFGKGVKFKALDKGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVD 2615

Db 2548 GHWFATTPIIGKIMFAIKEGRVTGVSIIASEDSRKVASVLNNAYYLDKMHYSIEGKD 2607

QY 2616 THYFVKPGPSEGDLAILGLSGRRTELVGNVTVSQINTVLNGRTTTRYTDIQLQYGALCL 2675

Db 2608 THYFVKIGAADGDLVLTGTTIGRKVLESGVNVTVSQPTLLVNGRTRRTFTNIEFOYSTLLL 2667

QY 2676 NTRYG---TTLDEEKARVLELARAVRAVQAWAREQQRLRGEGLRAWTEGEKQOVLSTG 2732

Db 2668 SIRYGLTPTLDEEKARVLDQAGRALGTAWAKEQQKARDGREGSRLTWTEGEKQQLSTG 2727

QY 2733 RVQGYDGFVVISVEQYPELSDSANNIHFMRQSEMGR 2769

Db 2728 RVQGYEGYVLPVEQYPELADSSNIQFLRQNMGR 2764

RESULT 9

Q9WTS4

ID Q9WTS4 PRELIMINARY; PRT; 2731 AA.

AC Q9WTS4;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Ten-m1.

GN ODZ1 OR TEN-M1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Balb/c;

RA Oohashi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,

RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;

RT "Mouse Ten-m/Odz is a new family of dimeric type II transmembrane

RT proteins expressed in many tissues.";

RL J. Cell Biol. 0:0-0(1999).

DR EMEL; AB025410; BAA77396.1; --

DR HSSP; P00750; ITPG.

DR MGD; MGI:1345185; Odz1.

DR GO; GO:0005887; C:integral to plasma membrane; IDA.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR001258; NHL.

DR InterPro; IPR000209; Peptidase\_S8.

DR InterPro; IPR006530; YD.

DR Pfam; PF00008; EGF; 5.

DR Pfam; PF01436; NHL; 5.

DR Pfam; PF05593; RHS repeat; 6.

DR SMART; SM00181; EGF; 7.

DR TIGRFAMS; TIGR01643; YD\_repeat\_2x; 7.

DR PROSITE; PS00022; EGF\_1; 8.

DR PROSITE; PS01186; EGF\_2; 7.

DR PROSITE; PS00136; SUBTILASE\_ASP; 1.

KW EGF-like domain.

SQ SEQUENCE 2731 AA; 305792 MW; 9129FA4CFE4A7770 CRC64;

Query Match 64.5%; Score 9607; DB 11; Length 2731;

Best Local Similarity 62.5%; Pred. No. 0;

Matches 1754; Conservative 407; Mismatches 533; Indels 112; Gaps 19;

QY 1 MDVKERKPYRSLTR-RRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGS- 58

Db 1 MEQTDCKPYQLSKVKHEMDLAYTSSDSEDEGKRPQSFNSRETLHEYNQELRRNYSQ 60

QY 59 --RVKDIVPQAE-EFCRTGANFTLRELGLEEVTPPHGTL--YRTDI-GLPQCGYSMGA 111

Db 61 SRKRKDVKEKSTQIEFCE-----TPP--TLCGYHTDMHSVSRHGYQLEM 103

QY 112 GSDADMEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQ 171

Db 104 GSDVDTETEGAAASPDHALRMWIRGMKSEHSSCLSSRANSALSLTDTDHER-KSDGNGFK 162

QY 172 -----NHAHLR--TPPPPLSHAHTPNQHAASINSNLRGNFT 206

Db 163 FSPVCCDMEAPADSAQDMQSSPHNQFTFRPLPPPPPPHACTCARKPPPTVDSLQRRSMT 222

QY 207 PRSNPSPAPTDSLSGEPAGGAQEPAAQENWLLNSNIPLETNLGKQPFGLQLDNLI 266

Db 223 TRSQSPA-----APAPPTSTQDSVHLHNSVWLSNIPLETN----- 259

QY 267 EMDILGASRHDGAYSDDHFLFKPG-GTSPLFCTTSPGYPLTSSVTYSPPPRPLRSTFAR 325

Db 260 -----HFLFKHSGSSAIFSAASQNYPLTSNTVYSPPPRPLRSTFSR 302

QY 326 PAFNLKPKSKYCNWKCAALSAIVISATLVLLAYFVAMHLFGLNWHLQPMQMEGYEITED 385

Db 303 PAFTFNKPYRCCNWKCTALSATAITVTALLLAYVIAVHLFGLTWQLQPV-GQIYANGIS 361

QY 386 TASSNEVPVPTDVSLYPSGGTGLETPDRKGGTTEGKPSFFPEDSFIDSGEIDVGERRASQK 445

Db 362 NGNPGETESMDTTYSPIGGRVSDKSEKK-----VFQKGRAIDTGEVDIGAQMVT 410

QY 446 IPPGTFRWSQVFIHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQTFDFVELLDGRRLLTQ 505

Db 411 IPPGLFWRFQITIHHPILYLFNLSLAKDSLGLIYGRNIPPTHTQTFDFVKLMDGKQLVKQ 470

QY 506 EARSLEGTPRQSRGTVPSPSHETGFIQYLDSGIWHLAFYNLDGKSEVVSFLTIAIESVDN 565

Db 471 DSKSSDDIQHSRNLILTSLQETGFIYMDQGPWLYAFYNDGKKMEQVFLTTAIEIMDD 530

QY 566 CPSNVCYNGDCISGTCHCFGLGFLGPDCCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECDV 625

Db 531 CSTNCGNGECISGHCHCFGLGPDCCARDSCPVLCCGNGEYKGHGVCVRNGWKGPCECDV 590

QY 626 PTNQCIDVACSNHGTCTGTCTICNPGYKGESCEEVDCMDPTCSGRGVCVVRGECHCFVGVWG 685

Db 591 PEEQCIDPTCFGHGTCIMGVCI CVPGYKGEICBEEDCLDPMCSSHGICVKGECHCSTGMG 650

QY 686 GTNCETPRATCLDQCSGHGTFLPDTGLCSCDSPSWTGHDCSIEICAADCGGHGVCVGTCR 745

Db 651 GVCNETPLPICQEQCSGHGTFLDGTVCSCDPKWTGSDCSTELCTMECASHGVCSRGICQ 710

QY 746 CEDGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCN 805

Db 711 CEEGWVGTCEERSCHSHCAEHGQCKDGKCECSPGWEGDHCTIAHYLD-AVRDGCPLCF 759

QY 806 GNGRCTLDLNGHWCVCQLGWRGAGCDTSMETACDSKDNDDGLVDCMDPDCCLOPLCHI 865

Db 770 GNGRCTLDQNGHWCVCQVGSWGTGCNIVMEMLCGDNLNDGDGLTDCVDPDCCQSQNCYV 829

QY 866 NPLCLGSPNPLDIIQETQVPVVSQQNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIR 925

Db 830 SPLCQGSPLDLIIQCSQPLFSQHTSRLFYDRIKFLIGKDSHTHVVPQDISFDSRRACVIR 889

QY 926 GQVMTSDGTPLGVNISFVNNPLFGYTIISRQDGSFDLVTNGGISIILRPERAPFITQEHT 985

Db 890 GQVAVDGTPLGVNVSFLHSDYGTIISRQDGSFDLVAIGGISVVLIFDRSPFLSEKRT 949

QY 986 LWLPWDRFFVMEETIIMRHEENEIIPSCDLNFAFPNPVPSPLTSFASCAEKGPIVPEI 1045

Db 950 LWLPWNQFIVVEKVINQRIVADAPSCDISNFI SPNPIVLPSPLTSTFGGSCP ERGTIVPEL 1009

QY 1046 QALQEEISISGCKMRLSYLSRTPGYKSVLRISLTHPTIPFNLMKVHLMVAVEGRLFRKW 1105

Db 1010 QVVQEEIPIPSFVRLSYLSRTPGYKTLRLTHSTIPVGMIKVHLTVSVEGRLTQKW 1069

QY 1106 FAAAADLSYFIWKTDVYNQKVFGLSEAFVSQVGEYESCPDLILWEKRTVLQGEIDA 1165

Db 1070 FPAAINLVYTFWANKTDIYQKVWGLAEALVSQVGEYEMCEPFIWEQRTVVLQGFEMDA 1129

QY 1166 SKLGGWSLDKHHALNIQSGTILHKGNGENQFVSQOQPVIGSIMGNGRRRSISCPSCNGLAD 1225

Db 1130 SNLGGWSLNKHHIFNPQSGIHKNGENMFISQOQPVVIATIMGNGHQRSVACTNCGPAH 1189

QY 1226 GNKLLAPVALTCGSDGSLYVGDFFNYIRIFPSPGNVTNILELRNKDFRHSHPAKYYLAT 1285

Db 1190 NNKLFAPVALASGPDGCVYVGDFFNVRRIFPSPGNSVSILELRNRDTRHSTSPAKYYLAM 1249

QY 1286 DPMGSAVFLSDSNSRRVPKIKSTVVVKDLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEA 1345

Db 1250 DPMSESLYLSDTNTRKVKYKLSLVETKDLKSNFEVVAGTGDQCLPFDDQSHCGDGGKASEA 1309

QY 1346 TLTNPRGITVDKFGLLYFVDGTMIRRIDONGIISTLLGSNDLTSARPLSCDSVMDISOVR 1405

Db 1310 SLNSPRGITVDRHGFYFVDGTMIRRIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVR 1369

QY 1406 LEWPTDLAINPMDNSLYVLDDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHA 1465

Db 1370 LEWPTDLAVNPMDSLYVLDDNNIVLQISENRRVRIIAGRPIHCQVPGIDHFLVSKVAIHS 1429

QY 1466 TLESATALAVSHNGVLYIAETDEKKINRIRQVTTSGEISLVAGAPSGCDCKNDANCDDFS 1525

Db 1430 TLESARAI SVSHSGLLFI AETDERKVNRIQQVTTNGEISIIAGAPTDCDCKIDPNCDFS 1489

QY 1526 GDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRIRPKNKPFLNTQNMVELSSPIDQE 1585

Db 1490 GDDGYAKDAKMAPSSLAVSPDGTLYVADLGNVIRITISKNAHLNDMNLVBIASPADQE 1549

QY 1586 LYLFDTTGKHLTQSLPTGPLYNFYTYTGDGDTILTDNNGNMVNVRDRSGMPLWLWVP 1645

Db 1550 LYQFTVNGTHLHTMNLITRDYVYNYFTYNAEGDLGALTSSNGNSVHIRRDAGGMPLWLWVP 1609

QY 1646 DGOVYVVTMTGNSALKSVTTQGHHELAMMTYHGNSGLLATKSNENGTWTTFFEYDSFGRLTN 1705

Db 1610 GGOVYVWLTISSNGVLKRVSAQYNLALMTYPGNTGLLATKSNENGTWTTVVEYDPEGHLTN 1669

QY 1706 VTPEPTQGVSSFRSDTSSVHVQVETSSKDDVTITTNLSASGAFYTLLOQDVNSYIIGAD 1765

Db 1670 ATPEPTGEVSSFHSDLEKLTKVALDTSNRENVLMSNTLATSTIYILKQENTQSTYRVSPD 1729

RESULT 10  
Q9R1K2

PRELIMINARY; PRT; 2765 AA.

ID Q9R1K2

QY 1766 GSLRLLLANGMEVALQTEPHLLAGTVNPTVGRNVTLPIDNGLNLVWRQKEQARGQVT 1825

Db 1730 GSLRVTFASGMEINLSSEPHILAGAVNPTLGKCNISLPGEHNANLIEWRQKEQKGNVS 1789

QY 1826 VFRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSRLNGVNV 1885

Db 1790 AFERRLRAHNRNLLSIDFDMTRTKIYDDHRKFTLRILYDQTRPILWSPVSRVNEVNI 1849

QY 1886 TYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFADKWTWSTYILEKSMVLLHSHQRYIF 1945

Db 1850 TYSPSGLVTFIQRGTWNERMEYDQSGKIIISRTWADKWTWSTYILEKSMVLLHSHQRYIF 1909

QY 1946 EFDKNDRLSSVTMPNVARQTLTIRSVGYRNIYQPEGNASVIQDFTEDGHLHFTPYLG 2005

Db 1910 EYDQSDCLLSVTMPSMVRHSLQTMLSVGYYRNIYTPDSDSTFIQDYSRDGRLLQTLHLG 1969

QY 2006 TGRVIYKYKGLSKLAETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQICPLID 2065

Db 1970 TGRVLYKYTKQARLSEILYDTTQVTLTYEESGVIKTIHLMHDGFICTIRYRQTGPLIG 2029

QY 2066 RQJFRFTEEGMVNARFDYVNSFRVTSMQAVINETPLIDLYRYDDVSGKTEQFGKFGV 2125

Db 2030 RQJFRFSEGLVNARFDYSY-NNFRVTSMQAVINETPLIDLYRYDDVSGRTEQFGKFSV 2088

QY 2126 IYDINQIITTAVMTHTKHFDAYGRMKEVQYEIFRSLMYMMTVQYDNMGRVVKKELKVGP 2185

Db 2089 INVDLNQVITTTVMKHTKIFNANGQVIEVQYEILKATAYWMTIQYDNMGRMVICDIRVGV 2148

QY 2186 YANTRYSYEYDADGQLOQTVSINDKPLWRYSDYDLNGLHLLSPGNSARLTPLRYDIRDRI 2245

Db 2149 DANTRYFEYDADGQLOQTVSVNDKIQWRYSYDLNGLNINLLSHGNSARLTPLRYDLRDRI 2208

QY 2246 TRLGDVQYKMWDEDFLQRGQDIFEYNSAGLLIKAYNAGSWSVRYRYDGLGRRVSSKSS 2305

Db 2209 TRLGEIQYKMWDEDFLQRGQNDIFEYNSNGLLQKAYNKVSGWTQYYYDGLGRRVASKSS 2268

QY 2306 HSHLQFFYADLTNPTKVTHLYNHSSSEITSLYYDLOHGLFAMELSSGDEFYIACDNIGT 2365

Db 2269 LGQHLQFFYADLANPIRVTHLYNHTSAEITSLYYDLOHGLIAMELSSGEEYYVACDNMGT 2328

QY 2366 PLAVPSGTGLMIKOILYAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRRDYVLA 2425

Db 2329 PLAVPSSRGQVIKEILYTPYGDYHDTYDPDFEVIIGFHGGLYDPLTKLVHLGQRDYDVVA 2388

QY 2426 GRWTPSDHELWKHLSSSNVM--PENLYMFKNNNPISNSQDIKCFMTDVNSWLLTTFGQLH 2483

Db 2389 GRWTPNHHIWQL--NLLPKPENLYSFENNYPVGKIQDVAKYTTDGTWLELFGQLH 2445

QY 2484 NVIPGYPKPMDAMEPSYELIHTQMKTOEWDNSKSILGYQCEVQKQKAFVTLERFDQLY 2543

Db 2446 NVLPGPFPKELENMELTYELLQQTKTQEWDPGKMILGIQCELOKQLRNFISLDQLPMT 2505

QY 2544 GSTITSCQAPKTKKFASSGSGVFGKVKFALKQGRVTTDIISVANEDGRRVAAAILNHAHY 2603

Db 2506 QYNEGRCLEGGKQPRFAAVPSVFGKIKFAIKEGIVTADIIGVANEDSRRLAAILNHAHY 2565

QY 2604 LENLHETIDGVDTHYFVKPGPSEGDLAAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRY 2663

Db 2566 LENLHETIEGRDTHYFIKLSLEEDLVLIGNTGGRRILENGVNVTVSQMTSVLNGRTRF 2625

QY 2664 TDIQLOYGALCLNTRYGTTLDDEKARVLELARQAVRQAWAREQORLREGEEGLRAWTEG 2723

Db 2626 ADIQLQHGALCFNIRYGTTVVEEKNHVLEMARQAVAQAWTQEQRRLOEGEGEGRVWTEG 2685

QY 2724 EKQVLSTGRVQYDGFVIVSVEQYPELSDSANNIHFMRQSEMGR 2769

Db 2686 EKQQLLGTGRVQYDGYFVLSVEQYLELSDSANNIHFMRQSEIGRR 2731

AC Q9RIK2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Neurestin alpha.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague Dawley; TISSUE=Olfactory bulb;  
RX MEDLINE=99350226; PubMed=10419693;  
RA Otaki J.M., Firestein S.;  
RT "Neurestin: putative transmembrane molecule implicated in neuronal development.";  
RL Dev. Biol. 212:165-181(1999).  
DR EMBL; AF086607; AAD47383.1; -.  
DR HSSP; P00750; ITPG.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001258; NHL.  
DR InterPro; IPR006530; YD.  
DR Pfam; PF01436; EGF; 4.  
DR Pfam; PF01436; NHL; 6.  
DR Pfam; PF05593; Rhs repeat; 5.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR SMART; SM00181; EGF; 5.  
DR TIGRFAMs; TIGR01643; YD\_repeat\_2x; 7.  
DR PROSITE; PS00022; EGF\_1; 8.  
DR PROSITE; PS01186; EGF\_2; 7.  
KW EGF-like domain.  
SQ SEQUENCE 2765 AA; 306568 MW; 6748D70D5FFD8F0E CRC64;  
  
Query Match 64.5%; Score 9601; DB 11; Length 2765;  
Best Local Similarity 61.8%; Pred. No. 0;  
Matches 1767; Conservative 422; Mismatches 488; Indels 180; Gaps 27;  
  
QY 1 MDVKERKPYRSLRRR-DAERRYTSSSADSEEGKAP-QKSYSSSETLKAYDQDARLAYGS 58  
DB 1 MDVKORR-HRSLTRGRCGKCRYTSSSLDSEDCRVPTQKSYSSSETLKAYDHSRMYGN 59  
  
QY 59 RVKDIVPQAEAEFCRTGANFTLRELGLEEVTPPHGTLRTDITGLPQCGYMGAGSDADWE 118  
DB 60 RVTDLVHRESDEFSRQGANFTLAEUGICEPS-PHRSGYCSDMGILHQYSLSTGSDADSD 118  
  
QY 119 ADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHENTETDHPG----- 168  
DB 119 TEGGMSPEHAIRLWGRGIKRRSSGLSSRENSALTTLTDSNENKSDDDNGRPIPTSSSS 178  
  
QY 169 -----G 169  
DB 179 LLPSAQLPSSHNPVPSVSCQPLLDSTNTSHQIMDTNPDEEFPNSYLLRACSGPQOASSG 238  
  
QY 170 LQNH---AELRTPPPPLSHAHTNQHHAASINSLRNGNETPRSN----PSPAPTDHSLSG 222  
DB 239 PPNHHSQSLRPPPLPP-PHNHILSHH-SSANSNRNSLTNRSQIHAPAPAND--LAT 294  
  
QY 223 EPPAGGAQEPHAQENWLLNSNIPLETNRNLGKQFFLGLQDNLIEMDILGASRHDGAYSD 282  
DB 295 TP-----ESVQLQDSWLVNSNVPLETR----- 316  
  
QY 283 GHFLFK-PGGTSPLFCTTSPGYPLTSSSTVYSPPPRPLRSTPARPAFNLKPKSKYCNWKC 341  
DB 317 -HFLFKTSSGSTPLFSSSSPGYPLTSGTVYTPPPRLLPRNTFSRKAFKLKPKSKYCSWK 375  
  
QY 342 AALSATVISATLVILLAYFVAMHLFGLNWHLQPMEGQMYEITEDTASSWPVPTDVSLYPS 401  
DB 376 AALSATAAALLAILLAYFIAMHLGLNWLQPADGHTF--NNGVRTGLPGNDVATVPS 433  
  
QY 402 GGTGLETPDRKGGKGTTEGKPSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHP 461

DB 434 GGGK-----VPWSL--KNSSIDSGEAEVGRRTVQEVPPGVFWRSQIHISQP 476  
QY 462 VHLKFNVSLGKAALVGIYGRKGLPPSHQTQDFVELLDGRRLLLTQBARSLGTPRQSRGTV 521  
DB 477 QFLKFNISLGKDALFGVYIRRGGLPPSHAQYDFMERLDGK-----EKWSVWESPRRRSIQ 531  
QY 522 PPSSETGTFIQYLDSGIWHLAFYNDGKSEVVSFLTTAIESVDNCPNSCYNGDCLISGTC 581  
DB 532 TLVQNEAVFQYLDVGLWHLAFYNDGKEMVSFNTVVLDVQDCPRNCHNGECVSGLC 591  
QY 582 HCFLGFLGPCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTIC 641  
DB 592 HCFPGFLGADCAKAAACPVLCSGNGQYSGKTCQCYSGWKGAECVPMNQCIDPSCGGHGS 651  
QY 642 ITGTCICNPGYKGESCEEVDMDPTCSGRGVCVRGECCHCFVGGTNCETPRATCLDQCS 701  
DB 652 IDGNCVCAAGYKGEHCEEVDCLDPTCSSHGVGVNGECLSPGWGGLNCELARVQCPDQCS 711  
QY 702 GHGTFPLPDTGLCSCDPSWTGHDCSIETCAADCGGHGVGVGGTCCEDGWMGAACDQACH 761  
DB 712 GHGTYLPDSGLCNCDPNWMGPDCEVSVCSVDCTGHGVCIGGACRCEEGWTGAACDQVCH 771  
QY 762 PRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDVVVKEGCPGLCNGNRCRTLDLNGMHCVC 821  
DB 772 PRCEIHGTCXDKGKCECREGWNHECTI-----DGCPCDLCNGNRCRTLQNSWQCVC 822  
QY 822 QLGWRGAGCDTSMETACGDSKNDGDGLVDMDPCDCLQLPLCHINPLCLGSPNPLDIQE 881  
DB 823 QTGWRGPGCVNVMETSCADNKNDEGLVDCLDPCDCLQSAQNSLLCRGSRDPLDIIQQ 882  
QY 882 TQVPVSQONLHSPYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOVMTSDGTPLVGNI 941  
DB 883 GQ--TDWPAVKSFYDRIKLLAGKDSHIIIPGDNPFNSLSLIRGOVTTDGTPLVGVNV 940  
QY 942 SFVNNPLFGYTISRQDGSFDLVTNGGISIILRFERAPFITQEHTLWLPWDRFFVMTIIM 1001  
DB 941 SFVKYPKYGYTITRQDGTDFDIANGCSALTLLHFERAPFMSRERTVWPPWNSFYAMDVLV 1000  
QY 1002 RHEENIPSCDLSNFARPNPVVSPSLTSPASSCAEKGIPIVEIQALQEEISISGCKMRL 1061  
DB 1001 KTEENSIPSCDLSGFVRPDPPIIISPLSTFFSAPANPIVPEPQVLHEEIELPGTNVKL 1060  
QY 1062 SYLSSRTPGYKSVLRISLTHPTIPFNLKMLVMAVEGRLPFRKFAAAPPDLSYFIWDKT 1121  
DB 1061 RYLSSRTAGYKSLKITMTQSTVPLNLRVHLMVAVEGLHFQKSFQASPNLAYTFIWDKT 1120  
QY 1122 DVYNQKVEGLSEAFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHALNI 1181  
DB 1121 DAYGQRYVGLSDAVSVGFYEYETCPSLILWEKRTALLQGFELDPNLSGGWSLDKHTLV 1180  
QY 1182 QSGILHKGNGENQFVSQPPVIGSIMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDG 1241  
DB 1181 KSGILLKGTGENQFLTQPPAIIITISIMGNGRRRSISCPSCNGLAEGNKLAPVALAVGIDG 1240  
QY 1242 SLVYVGDENYIRRIFFPSGNVTNILELRNKDFRHSHPAHKYVYLATDPMGAVFLSDNSRR 1301  
DB 1241 SLFVGDENYIRRIFFPSNRVNTSILELRNKEFKHSNPGHKYVYLAVDPTVGLSVSDTNSR 1300  
QY 1302 VFKIKSTVVVKDLVKNSEVAVGTGDCQLPFDTRCGDGGKATEATLTPRGITVDFKGLI 1361  
DB 1301 IYRVKSLSGAKDLAGNSEVAVGTGECQLPFDEARCGDGGKAVDATLMSPRGIAVDKNGLM 1360  
QY 1362 YFVDTGMIRRIDQNGIISTILGSDNLTSAARPLSCDSVMDISQVRLEWPTDLAINPMDNSL 1421  
DB 1361 YFVDATMIRKVDQNGIISTILGSDNLTAVRPLSCDSDMDVAQVRLEWPTDLAVNPMNSL 1420  
QY 1422 YVLNNVVLQISENHQVRIVAGRPMHQVPGIDHDFLLSKVAIHATLESATALAVSHNGVL 1481  
DB 1421 YVLENNVILRITENHQVSIAGRPMHQVPGID-YLSKLAIHSALESASAIASHTGVL 1479  
QY 1482 YIAETDEKKINRIQVTTSTSEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLTNPSS 1541

Db 1480 YITETDEKKINRLRQVTTNGEICLLAGAAASDCCKNDVNCICYSGDDAYATDAILNSPSS 1539

Qy 1542 LAVCADGELYVADLGNIRIRFIRKNKPFNTQNMYELSSPIDQELYLFDFTGKHLYTQSL 1601

Db 1540 LAVAPDGTIYIADLGNIRIRAVSKNKPVLNAFNQYEAASPEQEELYVFNADGIHQYTVSL 1599

Qy 1602 PTGDYLYNFTYTGDDITLITDNNGMVNVRDSTQMPLWLVPDQGVYVWMTGNTSALK 1661

Db 1600 VTGEYLYNFTYSADNDVTGLDNGNSLKIIRRDSSGPRHLLMPDNQIITLTVGTNGGLK 1659

Qy 1662 SVTTQGHELAMTYHGNSGLLTKSNENGWTTTFEYDYSFGRLTNVTFTPGQVSSFRSDTD 1721

Db 1660 AVSTQNLELGLMTYDGNLTGLLTKSDETGWTTTFYDYDHEGRLTNVTRPTGVVTSLHREME 1719

Qy 1722 SSVHVQVETSSKD-DVTIITNLSASGAFYTLQLQDVNSYYIGADGSLRLILLANGMEVAL 1780

Db 1720 KSITVDIENSRDNDVTITNLSSEVEASYTVVQDVNSYQLCSNGTLRVMYANGMVSF 1779

Qy 1781 QTEPHLLAGTVNPTVGKENVTLPIDNGLNLVEMRQKEQARGQVTVFGRRLRVHNRNLLS 1840

Db 1780 HSEPHVLAGTLTPTIGRCNISLPMENGLNSIEWRLRKEQIKGVITIFGRKLRVHGRNLLS 1839

Qy 1841 LDPDRVTRTEKIYDDHRKFTLIRILYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQRGI 1900

Db 1840 IDYDRNIRTEKIYDDHRKFTLIRIYDQGRPFLLWLPSSGLAANVSVYFFNGRLAGLQGA 1899

Qy 1901 MSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLSVVTMPN 1960

Db 1900 MSERTDIDKQGRIVSRMFADGKWSYSYLDKSMVLLLSQRQYIFEYDSSDRDLHAVTMP 1959

Qy 1961 VARQTLLETIRSVGYRNIYQPPPEGNASVIOQDFTEDGHLHLHTFYLGTGRRVIYKYGKLSKL 2020

Db 1960 VARHSMSTHTSIGYIRNIYNPPESNASVIFDYSDDGRILKTSFLGTGRQVYKYGKLSKL 2019

Qy 2021 AETLYDTTKVSFTYDETAGMLKTNLQNEGFTCTIRYQIGPLIDRQIRPRTTEEGMVNAR 2080

Db 2020 SEIYVDSSTAVTFGYDETTGVLLKMNVLQSGGFSCTIRYKVGPLVDKQIYRPFSEEGMINAR 2079

Qy 2081 FDYNY-DNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAVM 2139

Db 2080 FDYTHDNSFRIASIKPVISETPLVPDLRYDEISGKVEHFGKFGVIYDINQIITTAVM 2139

Qy 2140 THTKHFDAYGRMKEVQYEIFRSLMYMTVOYDNMGRVVKELKVGYPYANTTRYSEYDAD 2199

Db 2140 TLSKHFDTHGRIKEVQYEMFRSLMYMTVOYDSMGRVIKRELKGPYANTKYTYDYDGD 2199

Qy 2200 GOLQTVSINDKPLWYSVDLNGNLHLLSPGNSARLTPLRYDIRITRLGVDQYKMDDEG 2259

Db 2200 GOLQSAVANDRPTWRYSDVNLGNLHLLNPGNSARLMLPLRYDLDRITRLGVDQYKIDDDG 2259

Qy 2260 FLQRGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGRRVSSKSSHSHLQFFYADLTN 2319

Db 2260 YLCQSGSDIFEYNSKGLLTRAYNKASGWSVQYRYDGVSRRASAKTNLGHILQYFYSDLHH 2319

Qy 2320 PTKVTHLYNHSSEITSLYYDLQGHLFAMELSSGDEFFYIACDNIGTPLAVFSGTGLMIKQ 2379

Db 2320 PTRITHVYNHSNSEITSLYYDLQGHLFAMESSSGEBYYVASDNTGTPLAVFSINGLMIKQ 2379

Qy 2380 ILYTAYGEIYMDTNPNFQIIGYHGGLYDPLTLKLVMGRRDYDVLAGRWTSPDHKLWKLH 2439

Db 2380 LQYTAIGEIIYDSDNPDQFQMVIGFHGGLYDPLTLKVHFTQRDYDVLAGRWTSPDYTMWRNV 2439

Qy 2440 SSSNVMPFNLYMFKNNPNISNQDIKCFMTDVNSWLLTTFGFQLHNVIPGYPKPDMDAMEP 2499

Db 2440 GKEPA-PFNLYMFKNNPNLSNELDLKNYVTDVKSWLVMFGFQLSNIIIPGFPRAKMYFVPP 2498

Qy 2500 SYELIHTQMKTOEWDNSKILGVQCEVQKLFVTLERFDQLYGSTITSQQAPKTKK- 2558

Db 2499 PYELSESQAS-----ENGQLITGVQQTTERHNNQAFALF-----GQVISKKLHAGIREKA 2548

Qy 2559 ---PASSGSVFGKGVKFAKLDGRVTTDIIISVANEDGRRVAAIILNHAHYLENLHFTIDGVD 2615

Db 2549 GHWFATTPIIGKIGIMFAIKGRVTTGVSSIASEDSRKVASVLNNAYILDKMYHSIEGKD 2608

Qy 2616 THYFVKPCPSEGDIAILGLSGRRTLENGVNVTVSQINTVLNGRTRRRYTDIQLQYGALCL 2675

Db 2609 THYFVKIGAADGLVTLGTTIGRKVLESGVNVTVSQPTLLVNGRTRRFTNIEFYQSTLLL 2668

Qy 2676 NTRYG---TTLDEEKARVLELARQRAVRQAWAREQQRLRBEEGELRAWTEGEKQOVLSTG 2732

Db 2669 SIRYGLTPDRTLDEEKARVLDQARQALGTAWAKEQQKARDGREGSRLWTEGEKQQLLSTG 2728

Qy 2733 RVQGDGFFVISVEQYPELSDSANNIHFMRQSEMGR 2769

Db 2729 RVQGYEGYVLPVEQYPELADSSSNIQFLRQNEGMGR 2765

RESULT 11

Q9UKZ4

ID Q9UKZ4 PRELIMINARY; PRT; 2725 AA.

AC Q9UKZ4;

DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)

DT 01-OCT-2003 (TReMBLrel. 25, last annotation update)

DE Tenascin-M1.

GN TNM1.

OS Hmo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20025749; PubMed=10556288;

RA Brandau O., Schuster V., Weiss M., Hellebrand H., Fink F.M.,

RA Kreczy A., Friedrich W., Strahm B., Niemeyer C., Belohradsky B.H.,

RA Meindl A.;

RT "Epstein-Barr virus-negative boys with non-Hodgkin lymphoma are

RT mutated in the SH2D1A gene, as are patients with X-linked

RT lymphoproliferative disease (XLP).";

RL Hum. Mol. Genet. 8:2407-2413(1999).

DR EMBL; AF100772; AAF04723.1; -.

DR HSSP; P00750; ITPG.

DR Genew; HGNC:8117; ODZ1.

DR GO; GO:0005576; C:extracellular; TAS.

DR GO; GO:0008201; F:heparin binding; TAS.

DR GO; GO:0006955; P:immune response; TAS.

DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.

DR GO; GO:0007399; P:neurogenesis; TAS.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR001258; NHL.

DR InterPro; IPR000209; Peptidase\_S8.

DR InterPro; IPR006530; YD.

DR Pfam; PF00008; EGF; 5.

DR Pfam; PF01436; NHL; 6.

DR Pfam; PF05593; RHS\_repeat; 6.

DR SMART; SM00181; EGF; 6.

DR TIGRFAMs; TIGR01643; YD\_repeat\_2x; 7.

DR PROSITE; PS00022; EGF\_1; 8.

DR PROSITE; PS01186; EGF\_2; 7.

DR PROSITE; PS00136; SUBTILASE\_ASP; 1.

KW EGF-like domain.

SQ SEQUENCE 2725 AA; 304892 MW; 6F383E64C1B2E1B1 CRC64;

Query Match

Best Local Similarity 63.8%; Score 9501; DB 4; Length 2725;

Matches 1737; Conservative 404; Mismatches 550; Indels 112; Gaps 17;

Qy 1 MDVKERKPYRSLTR-RRDAERYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGS- 58

Db 1 MEQTDCKPYQPLPKVKHEMDLAYTSSDSESDGKPKRQSYNSRETLHEYNQELRMVNSQ 60

Qy 59 --RVKDIVPQEA-EFCRTGANFTLRELGLEEVTTPPHGTLYRTDI-GLPQCYSMGAGSD 114

Db 61 SRKRKEVEKSTQEMEFCEFTSHTLC-----SCYQTDHMSVSRHGYQLEMGSD 106

QY	115	ADMEADTVLSPHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEHENTETDHPGGLQ---	171
Db	107	VDTETEGAASPDHALRMWIRGMKSEHSSCLSSRANSALSJTDTDHER-KSDGENGFKEFSP	165
QY	172	-----NHAHLR--TPPPPLSHAHTPNQHEAASINSLNRGNFTPRS	209
Db	166	VCCDMEAQAGSTQDVQSSPHNQFTFRPLPPPPPPHACTCARKEPPPAADSLQRRSMITRS	225
QY	210	NPSAPTDSLGSPPAGGAQEPAAHQENWLLNSNIPLETRNLGKQPFLGTLQDNLIEMD	269
Db	226	QPSA-----APAPPTSTQDSVHLHNSWVLSNIPLETR-----	259
QY	270	ILGASRHDGAYSDGHFLFKPG-GTSPFLCTTSPGYPLTSSVTYSPPPRPLRSTFARPAF	328
Db	260	-----HSLFKHSGSSAIFSAASQNYPLTNTVYSPPPRPLRSTFSRPAF	305
QY	329	NLKPKSKYCNWKAALSAIVISATLIVLLAYFVAMHLFGLNWLHPMEGQMYEITEDTAS	388
Db	306	TENKPYRCCNWKCTALSATAITVTLALLLAYVIAVHLFGLTWIQLQVVEGELYANGVSKGN	365
QY	389	SWPVPTDVSLYPSGGTGLETPDRKGKGTGTEKPSFFPEDSFIDSGEIDVGRRASQKIPP	448
Db	366	RGTESMDTTYSPGGKVS DKSEK-----VFQKGRAIDTGEVDIGAQMOTIPP	414
QY	449	GTFWRSQVFIHDPVHLKFNVSLGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEAR	508
Db	415	GLFWRFQITIIHPIYLFKNISLAKOSLLGIYGRNIPPTHTQDFVVKLMDGKQLVKQDSK	474
QY	509	SLEGTPTQSRGTVPSPSSHETGFIQYLDGSIWHLAFYNDGKESSEVVSLTTAIESVDNCP	568
Db	475	GSDDTQHSPRNILITLSQETGFIEYMDQGPWYLAFYNDGKKMEQVFLTTAIEIMDDCST	534
QY	569	NCYNGDCISGTCHEFLGLPDPCCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECVPTN	628
Db	535	NCNGNCECISGHCHCFPGFLGPDCARDSCPVLCGGNGEYKHCVCRRHWKGPCECDVPE	594
QY	629	QCIDVACSNHGTCTGTICNPGYKGESCHEVD CMDPTCSGRGVCVGRGECHCFVGMGGTN	688
Db	595	QCIDPTCFGHGTCIMGVICVPGYKGEICBEEEDCLDPMCNSHGI CVKGECHCSTGWGGVN	654
QY	689	CETPRATCLDQCSGHGTFLPDTGLCSDPSWTGHDCSIEICAADCGGHGVCVGGTCRCHD	748
Db	655	CETPLPVCQECCSGHGTFLLDAGVCSDDPKWTGSDCSTELCTMECGSHGVCSRGICQCEE	714
QY	749	GWMGAACDQACHPRCAEHGTCDGKCECSPGWNGEHCTIAHYLDVWVKEGCPGLCNGNG	808
Db	715	GWVGPTCEERSCHSHCTEHGQCKDGKCECSPGWEGDHCTIAHYLD-AVRDGCPLCFNG	773
QY	809	RCTLDLNGWHCVQOLGWRGAGCDTSMETACGDSKONDGDLVDCMDPDCCLQPLCHINPL	868
Db	774	RCTLDQNGWHCVQVGWSGTGCNVVMMELCGDNLNDGDLTDCVDPCCQSQNCYISPL	833
QY	869	CLGSPNPLDIIQETQVPVSQQNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOV	928
Db	834	COGSPDPLDIIQSQTLFSQHTSRLFYDRIKFLIGKDSHTHVIPEVSFDSRRACVIRGOV	893
QY	929	MTSDGTPLVGVNLSFVNPLFGYTISRQGSFDLVNTGGISIIILRPERAPFITQEHTLWL	988
Db	894	VAIDGTPLVGVNVSLHHSYDGTISRQGSFDLVAGGISVILIFDRSPFLPEKRTLWL	953
QY	989	PWDRFFVMEIIMRHEENEIPSCDLSNFARPNPVSPSPLTSEASSCAEKGPVPEIQAL	1048
Db	954	PWQFIVVEKVTMQRVVDPPSCDISNFISPNPVLPSPLTSFGGSCPERGTIVPELOV	1013
QY	1049	QEEISIGCKMRLSYLSSRTPGYKSVLRISLTHPTIPFNLKMKVHLMVAVEGRLFRKWEAA	1108
Db	1014	QEEIPIPSFVRLSYLSSRTPGYKTLRLILTHSTIPVGMIVHLTVAVEGRLTQKWEPA	1073
QY	1109	APDLSYFIWKTQVYNQKVFGLSEAFVSQVGEYESCPDLILWEKRTTVLQGEIDASKL	1168
Db	1074	AINLVYTFANWKTQDIYGQVWGLAEALVSQVGEYETCPDFILWEQRTVVVLQGFEMDASNL	1133
QY	1169	GGWSLDKXHALNIQSILHKGNGENQFVSQQPPVIGSIMGNRRRSISCPSCNGLADGNK	1228

Db	1134	GDWSLNKXHLNPQSGIIHKGNGENMFISQPPVISTIMGNHGORSVACTNCGPAHNK	1193
QY	1229	LLAPVALTCGSDGSLYVGFDFNYIRIRIPPSGNVTNILELRNKDFRHSHPAKKYIATDPM	1288
Db	1194	LFAPVALASGPDGSVYVGFDFNFVRRIIPPSGNSVSILEL-----STSPAKKYIAMDVP	1246
QY	1289	SGAVFLSDNSRRVFKIKSTVVVVDLVKNSEVVAGTGDQCLPDDTRCGDGGKATEATLT	1348
Db	1247	SESLYLSDTNRKVYKLSLVETKDLKSNFEVVAGTGDQCLPDDQSHCGDGGRASEASLN	1306
QY	1349	NPRGITVDKFLYFVDGTMIRRIDQNGIISTLLGSNDLTSAPPLSCDSVMDISOVRLEW	1408
Db	1307	SPRGITVDRHGFYFVDGTMIRKIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVRLEW	1366
QY	1409	PTDLAINPMDNSLYVLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLE	1468
Db	1367	PTDLAVNPMDNSLYVLDNNIVLQISENRRVRIIAGRPIHCQVPGIDHFLVSKVAIHSTLE	1426
QY	1469	SATALAVSHNGVLYIAETDEKINIRQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDD	1528
Db	1427	SARAISSHSGLLFIAETDERKVNRIQQVTTNGEYIIAGAPTDCDKIDPNCDCFSGDG	1486
QY	1529	GYAKDAKLNTPSSLAVCADGELYVADLGNIRIRIRKKNPFLNTQNMVELSSPIDQELYL	1588
Db	1487	GYAKDAKMAPSSLAVSPDGTLYVADLGNVRIRITSRNQAHLNDMNIYEIASPADQELQ	1546
QY	1589	FDTTGKHLYTQSLPTGDLYNFTYTGDGDIITLITDNNGNMNVNRDSTGMPLWLVPDQG	1648
Db	1547	FTVNGTHLTLNLITRDYVYNFTYNSEGLGAIITSSNGNSVHIRRDAGGMPLWLVPDGG	1606
QY	1649	VYVWTMGTSALKSVTTQGHELAMTYHSGSGLLATKSNENGTTFYEYDSFGRLTNVTF	1708
Db	1607	VYWLTISSNGVLKRVSAQGYNPALMTYPGNTGLLATKSNENGTTVVEYDPEGHLTNATF	1666
QY	1709	PTQVSSFRSDTSSVHVQVETSSKDDVTITTNLSASGAFYTLQDQVRNSYYIGADGSL	1768
Db	1667	PTGEVSSFHSDLEKLTKEVLDTSNRNENVLMTNLTTATSTIYLKQENTQSTYRVPDGL	1726
QY	1769	RLLIANGMEVALQTEPHLLAGTVNPTVGKENVTLPIDNGLNLVWRQRKEQARGQVTVFG	1828
Db	1727	RVTFASGMEIGLSSEPHILAGAVNPTLGKCNISLPGEHNNANLIEWRQRKEQKGNVSAFE	1786
QY	1829	RRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVYS	1888
Db	1787	RRLRAHNRNLLSIDFDHITRTGKIYDDHRKFTLRILYDQTRPILWSPVSRYNENITYS	1846
QY	1889	PGYIAGIQRGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEED	1948
Db	1847	PSGLVTFIORGTWNEKMEYDQSGKIIISRTWADGKIWSYTYLEKSMVLLHLSQRRIIFEYD	1906
QY	1949	KNDRSSVTMPNVARQTLETIRSVGYRNIYQPEGNASVIQDPTEDGHLHLLHTFYLGTR	2008
Db	1907	QSDCLLSVTMPSMVRHSLQTMLSVGYYRNIYTPPDSSTSFIDYSRDGRLLQTLHLGTGR	1966
QY	2009	RVIYKYGKLSKLAETLYDTTKVSFTYDETAGMLKTINLQNEGTCTIRYQIGPLIDRQI	2068
Db	1967	RVLYKYTKQARLSEVLYDTTQVTLTYEBSGVIKTIHLMDHGFICTIRYQRTGPLIGRQI	2026
QY	2069	FRFTEEGMVNARFDYNDNSFRVTSQAVINETPLPIDLYDDVSGKTEQFGKFGVIYY	2128
Db	2027	FRFSEGLVNARFDYSY-NNFRVTSQAVINETPLPIDLYRYVDVSGRTEQFGKFSVINY	2085
QY	2129	DINQIITAVMTHTKHFDAYGRMKEVQYEIFRSLMYMTVQYDNMGRVWKELKVGYPAN	2188
Db	2086	DLNQVITTTVMKHTKIFSANGQVIEVQYELKAIAYWMTIQYDNVGRHGNMCIRVGVDAN	2145
QY	2189	TTRYSEYDADQLOQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDIRDRITRL	2248
Db	2146	ITRYFYEYDADQLOQTVSVNDKTQWRYSYDLNGDINLLSHGKSARLTPLRYDLRDRITRL	2205
QY	2249	GDVQYKMEDEGFLRQGGDIIFEYNSAGLLIKAYNRAGSWSRYRYDGLGRRVSSKSSHSH	2308

Db 2206 GBIQYKMWEDGFLRQRGNDIFEYNSNGLLQKAYNKASGWTQYYDGLGRRVASKSSLGQ 2265

QY 2309 HLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQGHLFAMELSSGDEFYIACDNIQTPLA 2368

Db 2266 HLQFFVDATANPIRVTHLYNHTSSEITSLYYDLQGHLIAMELSSGEEYYVACDNTGTPLA 2325

QY 2369 VFSGTGLMIKQILYTAIGEYIMDTNPNFQIIIGYHGGLYDPLTKLVHMGRDDYDVLAGR 2428

Db 2326 VFSSRGQVIKEILYTPGDIYHDTYDPFQVIIGHGGLYDFTLTKLVHLGQRDYDVVAGR 2385

QY 2429 TSPDHELKHLSSSNVM--PFNLYMFKXNNPNISNQDIKCFMTDVNSWLLTFGFQHLNVI 2486

Db 2386 TTAYHHIWKQL--NLLPKPFNLYSFENNYPVGKIQDVAKYTTDIRSWLELFGFQHLNVL 2442

QY 2487 PGYPKPDMDAMEPSYELIHTQMKTQEWDNKSILGVQCEVQKQKAFVTLERFDQLYGST 2546

Db 2443 PGFPKPELENLELTYYELLRLQTKTQEWDPGKTIIGIQCELQKQLRNFISLDQLPMTPRYN 2502

QY 2547 ITSCQAPKTKKFASSGSVFGKGVKFALKDGRVTTDIISVANEDGRRVAAILNHAHYLEN 2606

Db 2503 DGRCLEGGKQRFRAAVPSVFGKGIKFAIKDGIIVTADIIGVANEDSRRLAAILNNAHYLEN 2562

QY 2607 LHFTIDGVDTHYFVKPGPSEGDLAAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDI 2666

Db 2563 LHFTIEGRDTHYFIKLSLEEDLVLIQNTGGRILENGVNVTVSQMTSLNGRTRRFADI 2622

QY 2667 QLQYGALCLNTRYGTTILDBEKARVLELARQAVTQAWAREQORLREGEELRAWTEGEKQ 2726

Db 2623 QLQHGALCFNIRYGTTVBEEKNHVLEIARQRAVAQAWTKQORRLOEGEGEIRAWTEGEKQ 2682

QY 2727 QVLSTGRVQYDGFVTSVEQYPELSDSANNIHEMRQSEMGR 2769

Db 2683 QLLSTGRVQYDGYFVLSVEQYLELSDSANNIHEMRQSEIGRR 2725

RESULT 12

Q9W7R4 PRELIMINARY; PRT; 2590 AA.

AC Q9W7R4;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Ten-m3.

GN ODZ3 OR TENM3 OR TEN-M3.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99425191; PubMed=10495292;

RA Mieda M., Kikuchi Y., Hirate Y., Aoki M., Okamoto H.;

RT "Compartmentalized expression of zebrafish ten-m3 and ten-m4,

RT homologues of the Drosophila tenm /odd Oz gene, in the central nervous

RT system.";

RL Mech. Dev. 87:223-227(1999).

DR EMBL; AB026979; BAA81892.1; --

DR HSSP; P10968; 7WGA.

DR ZFIN; ZDB-GENE-990714-19; odz3.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR InterPro; IPR006209; EGF like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR002049; Laminin\_EGF.

DR InterPro; IPR001258; NHL.

DR InterPro; IPR006530; YD.

DR Pfam; PF00008; EGF; 3.

DR Pfam; PF01436; NHL; 5.

DR Pfam; PF05593; RHS repeat; 6.

DR PRINTS; PRO0011; EGFLAMININ.

DR SMART; SM00181; EGF; 6.

DR TIGRFAMS; TIGR01643; YD\_repeat\_2x; 7.

DR PROSITE; PS00022; EGF\_1; 8.

DR PROSITE; PS01186; EGF\_2; 7.

KW EGF-like domain.

SQ SEQUENCE 2590 AA; 288586 MW; 597592866213148D CRC64;

Query Match 62.5%; Score 9303; DB 13; Length 2590;

Best Local Similarity 63.4%; Pred. No. 0;

Matches 1685; Conservative 384; Mismatches 466; Indels 124; Gaps 19;

QY 140 RSSCLSSRANSNLTLTDEHENTE-----TDHPGGLQNHARLRTPPPLS 184

Db 27 QSSILSNATTQAVQDSDEEYTAVLYRPVTQPAPSHSCNEQPSNQHQGQSTLPPVPPP 86

QY 185 HAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGAQ-----EPAHAQENW 239

Db 87 HKQQP-----SVTALNHNLSLSSRRNVSPAP-----PAALPAELQTTPEVPLQDSW 132

QY 240 LLNSNIPLERNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPG-GTSPFLFCT 298

Db 133 VLGSNVPLESR-----HFLFKTGTGTPPLFST 159

QY 299 TSPGYPLTSSSTVYSPPPRPLPRSTFARPAFNLKKPSKYNWKCAALSIVAIVISATLVILLA 358

Db 160 ATPGYMATGAVYSPPTPLPRLNTLSRSFAFKFKSKXCSWRCTALSAMAVSILLVLLC 219

QY 359 YFVAMHLFGLNWHLQPMEGQMYE---ITEDTASSWPVPTDVSLYPSGGTGLETPDRKGK 415

Db 220 YCIAMHLFGLNWLQOETEGYAFENGQVKS DSTATNAV-----TALSTENK---- 264

QY 416 TTEGKPSFPPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFI DHPVHLKFNVSLGKAAL 475

Db 265 -----VYFQENNTIDTGEVDVGRRAVDVPPGTFWRTQLFIDQPQSLKFNISVQRGAL 317

QY 476 VGIYGRKGLPPSHTQTFDFVELLDGRRLLTQEARSL---EGTPRQSRGTVPPSSHETGFIQ 532

Db 318 VGVYGRKGLPPTHQYDFVELLDGSRLIAKEKRGLEVEGAARKARSV---NVHEABFIR 374

QY 533 YLDSGIWHLAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCCI SGTCFLGLGPD 592

Db 375 FLDSGTWHLAFYNDGKNAEQVSNTIITLTTECPHNCHNGDCRTGTCHCFPGFLGPD 434

QY 593 GRASCPVLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCITGTCINPGY 652

Db 435 SRAACPVLCSGNGQYSRGRCLCYSGWKGTCDVPSNQCIDIHCSGHGICIMGTCACNTGY 494

QY 653 KGESCEEVDCMDPTCSGRGVGVVGECHCFVGGTNCETPRATCLDQC SGHGTFLPDTGL 712

Db 495 KGDNCEEVDCLDPSCSSHGVCTHGECHCNPGWGGNNCEILKTMCPDQC SGHGTGYQTESGT 554

QY 713 CSCDPSWTGHDCSIEICAADCGGHGVGVGTCRCEDGWNMAACDQ RACHPRCAEHGTORD 772

Db 555 CTCDTNWTGPDCSI EVCAVDCSGSHGVCIGGSCRCEE GTSGVCDLKACHPRCTEHGTCKD 614

QY 773 GKCECSPGWNGEHCTIAHYLDRVVVKEGCPGLCNGNRCRTLDLNGWHVCVQLGWRGAGDT 832

Db 615 GKCECHQGTGEHCTV-----EGCPGLCNSNGRCTL DQNGWHVCVQPGWRGAGCDV 665

QY 833 SMETACGDSKNDGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQQNLH 892

Db 666 AMEILCADGKDSEGDGLVDCMDPDCCLOSSCQTQPFCEGSPDPFDIISQNQPASPOQAAQ 725

QY 893 SFYDRIKFLVGRDSTHIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFVNNPLFGYT 952

Db 726 SFYQQISFLTGPSTHVINGENPFNRSLVSIIRGQVLTDAGTPLIGVNVSVFVHPDHGYT 785

QY 953 ISRQDCGFDLVNTGGISIILRFERAPFITQHTLWLPDRFFVMEITIIMRHEENEIPSCD 1012

Db 786 ITQQQGMFDILANGGASLTLSFERAPFLTQFRVTWIPMNVFVYVMDTLVMKKEENDIPSCD 845

QY 1013 LSNFARPNPVVSPSLTISFASCAEKGPVPEIQALQEEIISIGCKMRLSYLSSRTPGYK 1072

Db 846 LSGFIRPSPLIVATPLSTFFRSSPENGPIIPETQVLQEETAIPGSDLNLMYLSRAAGYR 905

QY 1073 SVLRISLTHPTIPFNLMKVHLMVAVEGRLFRKWFAAAPDLSYFFIWDKTDVYNQKVFGLS 1132

Db 906 PVLKVTMTQATIPNLMKVHLMVAVVGRLPQKWFPAEPNLSYFIWDKTDAYNQRVYGLS 965  
QY 1133 EAFVSVGYEYESCDLILWEKRTTVLQGYEIDASKLGGWSLDKXHALNIQSILHKNGCE 1192  
Db 966 EAVSVGFYESCLDILWEKRTAILQGYELDASNMGWTLDXHVLVDVQNGILYKNGCE 1025  
QY 1193 NQFVSQOPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIR 1252  
Db 1026 NQFVSQOPPVISIMGNRRRSISCPSCNGQADGNKLLAPVALACGSDGSLFVGDFNYIR 1085  
QY 1253 RIFPSGNVTNILELNKDFRHSHPAHKYIYLATDPMGSAVFLSDSNRRVFKIKSTVVVK 1312  
Db 1086 RIFPSGNVTSVNEL-----SNNPAHGYYLATDPMTGQLYVSDTNSRRIFRPKALTGTK 1138  
QY 1313 DLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEATLTPRGITVDKFGLIYFVGDGTMIRRI 1372  
Db 1139 ELLQNAEVAAGTGEQCLPFDEARCGDGGKATEALLGPKGIAVDKNGFIYFVGDGTMIRKV 1198  
QY 1373 DQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDDNNVVLQI 1432  
Db 1199 DRNGIISTLLGSNDLTSARPLTCDNSMHIGQVRLEWPTDLAINPMDNSIYVLDDNNVVLQI 1258  
QY 1433 SENHQVRIVAGRMHCQVPGIDHFLSKVAIHATLESATALAVSHNGVLYIAETDEKKIN 1492  
Db 1259 TENQVRIVAGRMHCQVPGIE-YTMGKRAIQTTLEGATAISLSYSGVLYIAETDEKKIN 1317  
QY 1493 RIRQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDDGYAKADAKLNTPPSSLAVACDGLYV 1552  
Db 1318 RIRQVSTDGEISHLGAPSDCDCKNDANCDCYQTDGQYAKDARLNAPSSLVVSPDGLYV 1377  
QY 1553 ADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFDTTGKHLTYTQSLPTGDLYLNFTY 1612  
Db 1378 ADLGNIRIRALRHNRPPQSGSLFEVAPASQELYVFDNSNGTHQYTMSLVTGDYKYNFSY 1437  
QY 1613 TGDGDIITLTDNNGNMVNVRRDSTGMPLWLVPDQVYVWVTMGNTSALKSVTTQGHBLAM 1672  
Db 1438 SNEDDVTAVTDSSGNTLRVRDRPNRMPVRIVAPDNQVILWTIGTNGGLKTLTAQGOBLVL 1497  
QY 1673 MTHGNSGLLATKSNENGMWTFYEYDSFGRLTNVTFTPGQVSSFRSDTSSVHVQVETSS 1732  
Db 1498 FTYHGNSGLLATKSIQIGWTFDYDSEGRLTNTVFTPTGVTITSLIGEMDRALTVDIETSG 1557  
QY 1733 K-DDVTITTNLSASGAFYTLQDQVENSYYIGADGSLRLLLANGMEVALQTEPHLLAGTV 1791  
Db 1558 RDDDVSITTNLSSIDSFYTLVQDQLNSYQVGYDNSMRVIVANGMDSHFQTEPHILAGAS 1617  
QY 1792 NPTVGRNVTLPIDNGLNLEVRQKEQARGQVTVFGRRRLRVHNRNLLSLDFDVRTRTEK 1851  
Db 1618 NPTVARRNMTLPGENGQNLVEWRFRKEQNRGKVVFGRKLVRNGRNLSSVDYDRSLRTEK 1677  
QY 1852 IYDDHRKFLRILLYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQRIMSERMEYDQAG 1911  
Db 1678 IYDDHRKFLKIVYDASGHPPTLWVPSSKLSMVNLTYSSTGQVTSLQRPPTTERVEYDSQG 1737  
QY 1912 RITSRIFADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLSSVTMPNVARQTLTIRS 1971  
Db 1738 RIVSRTFADAKIWSYTYLDKSMVLLHLSQRQYIFDYDLHGKQIAITMPSVARHTMQTIRS 1797  
QY 1972 VGYRNIYQPPEGNASVIQDFTEDGHLHFTYLGTRRVYIKYKLSKLAETLYDTTKVS 2031  
Db 1798 VGYRNIYNPPESNASVTVDYSEDGQLLRVAHLGTGRRVLYKYRRQNKLSIILYDSTRVS 1857  
QY 2032 FTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVNARFDYNDNSFRV 2091  
Db 1858 FTYDETAGVLKTVNLQSEGFICIRYQIGPLVDRQIFRSESDGMVNARFDYTYDNSFRV 1917  
QY 2092 TSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITTAVMTHTKHFDAYGRM 2151  
Db 1918 TSMQGVINETPLPIDLYQFDDISGKVEQFGKFGVIYDINQIISTAVMTYTKHFDVHGRI 1977  
QY 2152 KEVQYEIFRSLMYWMTVQYDNMGRVVKELKVGYPYANTTRYSEYDADGQLOTVSINDKP 2211

Db 1978 KEIQYEIFRSLMYWITIQYDNMGRVTKRBIKIGPFANTTKYGYEYDVGQLOTVYLNEKM 2037  
QY 2212 LWRYSYDLNGLNHLHLLSPGNSARLTPLRYDIRDRITRLGQVQYKXNDEDDGFLRQRGDIFEY 2271  
Db 2038 MWRYNYDLNGLNHLNPGNSARLTPLRYDLRDRITRLGQVQYRDNDEDDGFLRQRGAEIFEY 2097  
QY 2272 NSAGLLIKAYNRAGSWVRYYRVDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSS 2331  
Db 2098 NSKGLLVRVHASKASGWTIQYRYDGLGRRRLASRNSLGHLQFFYADLNPTRITHVYNHSS 2157  
QY 2332 SEITSLYYDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYAYGEIYMD 2391  
Db 2158 SEITSLYYDLQHLFAMEISSGEEFYIACDNTGTPLAVFSSNGLLLKQVQYAYGEIYFD 2217  
QY 2392 TPNPFIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHLMKHLSSSNVMPFNLYM 2451  
Db 2218 SNPDLFQLVGFHGGLYDPLTRLHFGGRDYDIQAGRWTTPDISTWTRV-GKDPAPFNLYM 2276  
QY 2452 FKNNNPISNSQDIKCFMTDVNSWLLTVPFQQLHNVIPIGPKPMDAMEPSYELIHTQMKTK 2511  
Db 2277 FRNNNPISKIHEVKEYVTDVNIWLVTGFLHNVIPIGFPPIKFDLTQPSLEM----RKSQ 2332  
QY 2512 EWDNSKILGVOCEVQKQKAFVTLERFDQLYGSTITSCQQAQPKTKKFASSGSVFGKGVK 2571  
Db 2333 LWDDLPSISGVQQEVMRQAKAFLSFERMPEIQLSRRRSRREXP-WLWFATVKSLLIGKVM 2391  
QY 2572 FAL-KDGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGDLA 2630  
Db 2392 LAITSKGVATNALNIANEDCIKVTVLNNAFYLEDLHFTVEGRDTHYFIKTSLPESDLG 2451  
QY 2631 ILGLSGGRITLNGVNTVVSQINTVLNGRTRRYTDIQLQYCALCLNTRYGTTLDDEEKARV 2690  
Db 2452 ALRLTSGRKSLENGVNTVVSQITTVNGRTRRFADVELQYALALAHVRYGMTLDDEEKARV 2511  
QY 2691 LELARQRAVRQAWAREQORLREGEGLRAWTEGEKQOVLSTGRVQGYDGFVVISVEQYPE 2750  
Db 2512 LEQARQALSSAWAREQORVRDGEGLRWTEGEKROLLSSGKVLGYDGYVLSVEQYPE 2571  
QY 2751 LSDSANNIHFMRQSEMGR 2769  
Db 2572 LADSANNVQFLRQSEIGR 2590

RESULT 13

Q9JLCL1  
ID Q9JLCL1 PRELIMINARY; PRT; 2346 AA.  
AC Q9JLCL1;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE OD23 (Fragment).  
GN OD23.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ben-Zur T., Feige E., Motro B., Wides R.;  
RT "The mammalian Odz gene family: Homologs of a Drosophila pair rule  
RT gene with expression implying distinct yet overlapping developmental  
RT roles.";  
RL Dev. Biol. 117:107-120 (2000).  
DR EMBL; AF195418; AAF28316.1; -.  
DR HSSP; P01492; 1VNB.  
DR MGD; MGI:1345183; Odz3.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR001258; NHL.  
DR InterPro; IPR006530; YD.  
DR Pfam; PF00008; EGF; 3.

DR	Pfam; PF01436; NHL; 6.	
DR	Pfam; PF05593; RHS repeat; 6.	
DR	PRINTS; PRO0011; EGFLAMININ.	
DR	SMART; SM00181; EGF; 5.	
DR	TIGRFAMS; TIGR01643; YD_repeat_2x; 6.	
DR	PROSITE; PS00022; EGF_1; 8.	
DR	PROSITE; PS01186; EGF_2; 7.	
KW	EGF-like domain.	
FT	NON_TER	1
SQ	SEQUENCE	2346 AA; 261699 MW; D761563C6AEDD735 CRC64;
Query Match		
Best Local Similarity 61.4%; Score 9145; DB 11; Length 2346;		
Matches 1643; Conservative 319; Mismatches 371; Indels 54; Gaps 9;		
QY	392 VPTDVSLYPSGGTGLETPDRKGKGTTEGKPSFFPDSFIDSGEIDVGRRASQKIPPGTF	451
Db	5 VPTNVS LPSG-----DNGKLGGFTHENNTIDSGELDIGRRAIQEVPPGIF	50
QY	452 WRSQVFIDHPVHLKENVSLGKAALVGIYGRKGLPPSHQTQDFVELLDGRRLLTQEARSL	511
Db	51 WRSQLFIDQPOFLKFNISLQXDALIGVYGRKGLPPSHQTQYDFVELLDGSRLIAREORNLV	110
QY	512 GTPRQSRGTVPSSSHETGFIQYLDSGIWHLAFYNDGKSESEVVSFLTTAIESVDNCPSN	571
Db	111 ESERAGRQARSVSLHEAGFIQYLDSGIWHLAFYNDGKNPEQVSFNTIVIESVVECP	170
QY	572 GNGDCISGTCFGLGFLGPDCCRASCPLVCSNGQYMKGRCLCHSGWKABCDVPTNQCI	631
Db	171 GNGECVSGTCHCFPGFLGPDCCSRAACPVLCSNGQYSGKRCCLCFSGWKGTEDVPTTQCI	230
QY	632 DVACSNHGTCITGTCINPGYKGESCEEVDCMDPTCSGRGVCVRGECHCFVGWGGTNC	691
Db	231 DPQCGGRGICIMGSCACNSGYKGENCEEADCLDPGCSNHGVCIIHGECHCNPGWGSN	290
QY	692 PRATCLDQCSGHGTFLPDTGLCSCDPSWTGHDCSIICAADCGHGVCVGTCRCEDG	751
Db	291 LKTMCADQCSGHGTYLQESGSCDTPNWTGPDCSNEICSVDCSGHGVCMGGSRC	350
QY	752 GAACDQRACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNG	811
Db	351 GPACNQRACHPRCAEHGTCRDGKCECSQGWNGEHCTI-----EGCPGLCNSGR	401
QY	812 LDLNGHWCVCQLGWRGAGDTSMETACGDSKNDGDGLVDCMDPDCCQLPLCHINPL	871
Db	402 LDQNGHWCVCQPGWRGAGCDVAMETLCTDSKDNEDGLIDCMDPDCCQLQSSCQNP	461
QY	872 SPNPLDIIQETQVPVSQQNLHSEFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIR	931
Db	462 LPDPQDIIISQSLQTPSQQAASFYDRISFLIGSDSTHVLPGESPFNKSLASVIR	521
QY	932 DGTPLGVNISFVNPLFGYTISRQDGSFDLVINGGISIILRFERAPFITQEHTLWLP	991
Db	522 DGTPLGVNVSFLHYSEYGYTITRODGMEDLVANGGASLTLVFERSPFLTQYHTV	581
QY	992 RFFVMEIIMRHEENIPSCDLSNFARPNPVVSPSLTSFASSCAEKGPIVPEIQALQ	1051
Db	582 VFYVMDTLVMKKEENDIPSCDLSGFPVRPSPIIVSSPLSTFFRSSPEDSPIIP	641
QY	1052 ISISGCKMRLSYLSRTPGYKSVLRISLTHPTTIPFNLMKVHLMVAVEGRLFRKW	1111
Db	642 TTIPGTDKLSYLSRRAAGYKSVLKITMTQAVIPFNLMKVHLMVAVVGRLFKW	701
QY	1112 LSYFPIWKTVDVYNQVFGLSEAFVSVGYESPCDILWEKRTTVLQGEIDASKLG	1171
Db	702 LAYTFIWDKTDAYNQVYGLSEAVSVGYEYESCLDRTLWEKRTAVLQGYELDAS	761
QY	1172 SLDKGHALNIQSGILHKNGENQFVSQQPPVIGSIMGNRRRSISPCSNGLADGNK	1231
Db	762 TLDKHHVLDVQNGILYKNGENQFISQQPPVVSIMGNRRRSISPCSNQADGNK	821
QY	1232 PVALTCGSDGLYVGDRNYIRRIFFPSGNVTNILELRNKDFRHSHPAHKYVLATD	1291
	PMGA	1953

Db	822 PVALACGIDGSLYVGDFNYVRRIFPSPGNVTSVLEL-----SSNPAHRYYLATDP	874
QY	1292 VFLSDSNSRRVFKIKSTVVVVDLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEATLT	1351
Db	875 LYVSDTNTRRIYRPKSLTGAKDLTKNAEVAAGTGEQCLPFDEARCGDGGKAVEAT	934
QY	1352 GITVDKFGLIYFVDGTMIRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVR	1411
Db	935 GMAIDKNGLIYFVDGTMIRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISQVR	994
QY	1412 LAINPMDNSLYVLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHAT	1471
Db	995 LAINPMDNSIYVLDNNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQT	1053
QY	1472 ALAVSHNGVLYIAETDEKKNIRQVTTSGEISLVAGAPSGCDCKNDANCDCFS	1531
Db	1054 AIAVSYSGLYITETDEKKNIRQVTTDGEISLVAGIPSECDCKNDANCDCYQSG	1113
QY	1532 KDAKLNTPSSLAVCADGELYVADLGNIRIRFIRKKNKPFNTQNMVELSSPIDQEL	1591
Db	1114 KDAKLNAPSSLAASPDGTLTIADLGNIRIRAVSKNKLPLNSMNFYEVASPTDQEL	1173
QY	1592 TGKHLTQSLPTGDYLYNFTYTGDDITLITDNNGMNVVRDSTGMPLWLVPDQGV	1651
Db	1174 NGTHQYTVSLVTDYLYNFSYNDNDVTAVDTSNGNTLRIRDPNRMPPRVVSPD	1233
QY	1652 VTMGTNSALKSVTTQGHELAMTYHGNGLLATKSNENGWTTTFEYDVSFGRLT	1711
Db	1234 LTIGTNGCLKSMTAQGLELVLFYHGNGLLATKSDTEIGWTTFFDYDSEGRLT	1293
QY	1712 QVSSFRSDTSSVHVQVETSSK-DDVTITTNLSASGAFVTLLODQVRNSYIIGA	1770
Db	1294 VVTNLHGDMDKAITVDIESSSREEDVSITSNLSSIDSFYTMVQDLRNSYQI	1353
QY	1771 LLANGMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEWRQKEQARGQ	1830
Db	1354 FYASGLDSHYQTEPHVLAGTANPTVAKRNMTLPGENGQNLVEWFRKEQAQ	1413
QY	1831 LRNVNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSPSSRLNGVNT	1890
Db	1414 LRNVGRNLLSVDFDRTTKTEKIYDDHRKFLLRDAYDTSGHPTLWLPSSKL	1473
QY	1891 GYIAGIQRGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHHSQRQ	1950
Db	1474 GQIASIQRGTTSEKVDYDSQGRIVSRVAFADGKTWSYTYLEKSMVLLHHSQ	1533
QY	1951 DRLSSVTMPNVARQTLLETIRSVGYRNIYOPPEGNASVIQDFTEDGHLHTF	2010
Db	1534 DRLSAITMPSVARHTMQTIRSIGYRNIYNPPESNASITDYNEEGLLQTAFL	1593
QY	2011 IYKYGKLSKLAETLYDTTKVSYFTYDETAGMLKTNLQNEGFTCTIRYQIGPL	2070
Db	1594 LFKYRRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYQIGPL	1653
QY	2071 FTEEGMVNARFDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGK	2130
Db	1654 FSEDGMVNARFDYSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQ	1713
QY	2131 NQIITTAVMTHTKHFDAYGRMKEVQYEI FRSLMYMTVQYDNMGRVVKELK	2190
Db	1714 NQIISTAVMTYTKHFDAGHRIKEIQYEI FRSLMYWITIYQDNMGRVTKREI	1773
QY	2191 RYSVEYDADGQLOQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYD	2250
Db	1774 KYAYEYDVGQLOQTVVLYNEKIMWRNYDLNGNLHLLNPFSSARLTPLRYD	1833
QY	2251 VOYKXDEDDGFLRQGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGRRVSS	2310
Db	1834 VQVRLDEDDGFLRQRGTEIFEYSSKGLLTRVYSKSGWTVIYRYDGLGRRV	1893
QY	2311 QFFYADLTNPTKVTHLYNHSSEITSLYYDLQGHLEFAMELSSGDEFYIACD	2370
Db	1894 QFFYADLTYPTRITHVYNHSSEITSLYYDLQGHLEFAMEISSGDEFYIASD	1953

QY	2371	SGTGLMIKQILYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLGRWTS	2433
Db	1954	SSNGLMLKQIQYTAYGEIYFDSNVDFQLVIGTHGGLYDPLTKLIHFGERDYDILAGRWT	2013
QY	2431	PDHELWKHLSSNVNPFNLYMFKNNNPISNSQDIKCFMTDVSWLLTFGFQHLNVIPGYP	2490
Db	2014	PDIEIWKRI-GKDPAPFNLYMFRNNPASKIHVDVKDYITDVSWLVTFGFHLHNAIPGFP	2072
QY	2491	KPMDAMEPSYELIHTQMKTQEWDNKSILGVQCEVQKQLKAFVTLERFDQLYGSTITSC	2550
Db	2073	VPKFDLTSPSYELV---KSQQWEDVPPIFGVQQQVARQAKAFSLGKMAEV-----	2120
QY	2551	QQAPKTKK-----FASSGSVFGKGVKFKALDKDGRVTTDIIISVANEDGRRVAAILNHAH	2602
Db	2121	-QVSRKAGAEQSWLWFATVKSLLIGKGVMLAVSQGRVQTNVLNIANEDCIKVAAVLNNAF	2179
QY	2603	YLENLHFTIDGVDTHYFVKPSPSEGLAILGLSGGRTLENGVNVTVSQINTVLNGRTRR	2662
Db	2180	YLENLHFTIEGKTHYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVVNGRTRR	2239
QY	2663	YTDIQLQYGALCNTRYGTTLDEEKARVLELARQRAVRQAWAREQQRLREGEEGLRAWTE	2722
Db	2240	FADVEMQFALALHVRYGMTLDEEKARILEQARQALARAWAREQQRVRDGEGEARLWTE	2299
QY	2723	GEKQVLSGTRVGQDGFVISVEQYPELSDSANNIHFMRQSEMGR	2769
Db	2300	GEKQVLSAGKVGQDGYVLSVEQYPELADSANNIQFLRQSEIGKR	2346

RESULT 14

Q9ULU2

ID	Q9ULU2	PRELIMINARY;	PRT;	2144 AA.
AC	Q9ULU2;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Hypothetical protein KIAA1127 (Fragment).			
GN	KIAA1127.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=20039618; PubMed=10574461;			
RA	Hirosawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;			
RT	"Characterization of cDNA clones selected by the Genemark analysis			
RT	from size-fractionated cDNA libraries from human brain.";			
RL	DNA Res. 6:329-336(1999).			
DR	EMBL; AB032953; BAA86441.2; -.			
DR	GO; GO:0005509; F:calcium ion binding; IEA.			
DR	GO; GO:0005198; F:structural molecule activity; IEA.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	InterPro; IPR006209; EGF_like.			
DR	InterPro; IPR006210; IEGF.			
DR	InterPro; IPR002049; Laminin_EGF.			
DR	InterPro; IPR001258; NHL.			
DR	InterPro; IPR006530; YD.			
DR	Pfam; PF00008; EGF; 4.			
DR	Pfam; PF01436; NHL; 6.			
DR	Pfam; PF05593; RHM repeat; 5.			
DR	PRINTS; PR00011; EGFLAMININ.			
DR	SMART; SM00181; EGF; 6.			
DR	SMART; SM00179; EGF_CA; 2.			
DR	TIGRFAMS; TIGR01643; YD_repeat_2x; 5.			
DR	PROSITE; PS00022; EGF_1; 6.			
DR	PROSITE; PS01186; EGF_2; 5.			
KW	Hypothetical protein.			
FT	NON_TER			
SEQUENCE	2144 AA;	238607 MW;	B193948001AE46B9	CRC64;

Db	1018	HLMPDNQIITLTGTNGGLKVSVSTQNLELGLMTYDNGTGLLATKSDETGWTTFYDYDHE	1077
QY	1701	GRLTNVTFTPTGQVSSFRSDTDSSVHVQVETSSK-DDVTITTNLSASGAFYTLQDQVRNS	1759
Db	1078	GRLTNVTRPTGVVTSLHREMEKSIITIDIENSNRDDDDVTVITNLSVEASYTVVQDQVRNS	1137
QY	1760	YYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVFWRQRKEQ	1819
Db	1138	YQLCNGNGLRVMYANGMGISFHPSEPHVLAGTITPTIGRCNISLPMENGLNSIEWRLRKEQ	1197
QY	1820	ARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSR	1879
Db	1198	IKGKVTIFGRKLRVHGRNLLSIDYDRNIRTEKIYDDHRKFTLRILYDQVGRPLWLPSGG	1257
QY	1880	LNGVNVITYSPGGYIAGIQGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHS	1939
Db	1258	LAAVNVSYFFNGRLAGLQRGAMSSERTDIDKQGRIVSRMFADGKWSYSYLDKSMVLLQS	1317
QY	1940	QROYIFEFDKNDRLSSVTMPNVARQTLTETIRSVGYGYRNIYQPPPEGNASVIQFTEDGHLL	1999
Db	1318	QROYIFEYDSSDRLLAVTMPSPVARHSMSTHTSIGYIRNIYNPPESNASVIFDYDDGRIL	1377
QY	2000	HTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYRQ	2059
Db	1378	KTSFLGTGRQVFYKYGKLSKLSIYDSTAVTFGYDETTGVLKMVNLQSGGFSCTIRYRK	1437
QY	2060	IGPLIDRQIFRFTTEGMVNARFDYNY-DNSFRVTSMQAVINETPLPIDLYRDDVSGKTE	2118
Db	1438	IGPLAVDKQIYRFSEBGMVNARFDYTYHDNSFRIASIKPVISETPLPVDLYRDEISGKVE	1497
QY	2119	QFGKFGVIYYDINQIITAVMTHTKHFDAYGRMEKVQYEIFRSLMYWMTVQYDNMGRVVK	2178
Db	1498	HFGKFGVIYYDINQIITAVMTLSKHFDTHGRKEVQYEMFRSLMYWMTVQYDSMGRVIK	1557
QY	2179	KELKVGPIYANTTRYSYEYDADQLOQTVSINDKPLWRYSYDLNGNLHLSPGNSARLTPLR	2238
Db	1558	RELKLGPIYANTTRYKYDYDGDGQLQSVAVANDRPTRYSYDLNGNLHLNPGNSVRLMPLR	1617
QY	2239	YDIRDRITRLGDVQYKMDDEDGFLRQGGDIFEYNSAGLLIKAYNRAGSWSVRYRYDGLGR	2298
Db	1618	YDLRDRITRLGDVQYKIDDDGYLCQRGSDIFEYNSKGLLTRAYNKASGWSVQYRYDGVGR	1677
QY	2299	RVSSKSSHHLQFPFYADLTNPTKVTHLYNHSSEITSLYYDLQHLFAMELSSGDEFYI	2358
Db	1678	RASYKTNLGHLQYFYSDLHNPTRITHVYNHSSEITSLYYDLQHLFAMESSSGEEYV	1737
QY	2359	ACDNIGTPLAVFSGTGLMIKQILYATAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGR	2418
Db	1738	ASDNTGTPLAVFSGINGPMIKQLQYATAYGEIYSDNPDPFQMVIGFHGGLYDPLTKLVHFTQ	1797
QY	2419	RDYDVLAGRWTSPDHELWKHLSSSNVMPFNLYMEKNNPISNSQDIKCFMTDVNSWLLTF	2478
Db	1798	RDYDVLAGRWTSPDYTMWKNVGKEPA-PFNLYMEKSNPNLSSSELDLKNYVTDVKSWLVMF	1856
QY	2479	GFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDNKSILGVQCEVQKQLKAFVTLER	2538
Db	1857	GFQLSNIIPGFPRAKMYFVPPPYELSESQAS----ENGQLITGVQQTTERHNQAFMALE-	1911
QY	2539	FDQLYGSTITSCQAPKTKK----FASGSGVFGKGVKFPALKDGRVTTDIIISVANEDGRRV	2594
Db	1912	-----GOVITKKLHASIREKAGHWFATTTPIIGKIMFAIKEGRVTTGVSSIASEDSRKV	1966
QY	2595	AAILNHAHYLENLHFTIDGVDTHTYFVKPGPSEGDIALILGLSGRRRTLENGVNVTVSQINT	2654
Db	1967	ASVLNNAYYLDKMHYSIEGKDTHTYFVKIGSADGDLVTLGTTIGRKVLESGVNVTVSQPTL	2026
QY	2655	VLNGRTRRYTDIQLQYCALCLNTRYG---TTLDEEKARVLELARQRAVRQAWAREQQRRLR	2711
Db	2027	LVNGRTRRFTNIEFOYSTLLLSIRYGLTPDTLDEEKARVLDQARQARALGTAWAKEQOKAR	2086
QY	2712	EGEGLRAWTEGEKQOVLSTGRVQYDGGFFVISVEOYPELSDSANNIHFMRQSEMGRR	2769

Db 2087 DREGSRLWTEGEKQQLLSTGRVQGYEGYVLPVEQYPDLADSSNIQFLRQNMGKR 2144

RESULT 15

Q80TD2 PRELIMINARY; PRT; 1828 AA.

AC Q80TD2;

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE MKIAA1455 protein (Fragment).

GN MKIAA1455.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=22579291; PubMed=12693553;

RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,

RA Nakajima D., Nagase T., Ohara O., Koga H.;

RT "Prediction of the coding sequences of mouse homologues of KIAA gene:

RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous

RT cDNAs identified by screening of terminal sequences of cDNA clones

RT randomly sampled from size-fractionated libraries.";

RL DNA Res. 10:35-48(2003).

DR EMBL; AKL22513; BAC65795.1; -.

DR InterPro; IPR001258; NHL.

DR InterPro; IPR006530; YD.

DR Pfam; PF01436; NHL; 6.

DR TIGRFAMs; TIGR01643; YD\_repeat\_2x; 5.

FT NON\_TER 1

SQ SEQUENCE 1828 AA; 206341 MW; 5C2AAB91DC3DED8C CRC64;

Query Match 47.4%; Score 7063; DB 11; Length 1828;

Best Local Similarity 69.8%; Pred. No. 0;

Matches 1287; Conservative 269; Mismatches 263; Indels 24; Gaps 6;

QY 936 LVGNISFVNNPLFGYTISRQDGSFDLVNTGGISILRFERAPFITQEHILWLPWDRFFV 995

Db 1 LIGVNSFLHYSEYGYTITRQDGMFDLVANGGASLTIVFERSPFLTQYHTVWIPWNVFV 60

QY 996 METIMRHEENEIPSCDLSNFARPNPVSPSLTSEASSCAEKGPIVPEIQALQBEISIS 1055

Db 61 MDLTLYNKKBEENDIPSCDLSGFVRPSPPIIVSSPLSTFFRSSPEDSPIIPETQVLHEETIP 120

QY 1056 GCKMRLSYLSSRTPGYKSVLRISLTHPTIPFNLMKVHLMVAVEGRLFRKWFAPADLSY 1115

Db 121 GTDLKLSYLSSRAAGYKSVLKITMTQAVIPFNLMKVHLMVAVVGRFLQKWFAPASNLAYT 180

QY 1116 FIWDXTDVYNQKVGFLSEAFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDK 1175

Db 181 FIWDXTDAYNQKVGLSEAVVSVGYEYESCLDLTLWEKRTAVLQGYELDASNMGWTLDK 240

QY 1176 HHALNTQSGILHKNGENQFVSQQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVAL 1235

Db 241 HHVLDVQNGILYKNGENQFVSQQPPVSSIMGNRRRSISCPSCNGQADGNKLLAPVAL 300

QY 1236 TCGSDGSLYVGDFNYIRRIEIPSGNVNTNILELRNKDPRHSHSPAHKYIYLATDPMSCAVFLS 1295

Db 301 ACGIDGSLYVGDFNYVRRRIEIPSGNVTSVLELRNKDPRHSSNPAHRYIYLATDPVTGDLYVS 360

QY 1296 DSNRRVFKIKSTVVVKLVKNSVAVAGTGDQCLPDDTRCGDGGKATEATLTNPRGITV 1355

Db 361 DTNTRRIYRPKSLTGAKDLTKNAEVAAGTGEQCLPDEARCGDGGKAVEATLMSPKGMAI 420

QY 1356 DKFGLYFVDGTMIRRIDQNGIISTLLGSNDLTSAPLSCDSVMDISQVRLEWFTDLAIN 1415

Db 421 DKNGLIYFVDGTMIRKVDQNGIISTLLGSNDLTSAPLTCDSMHISQVRLEWFTDLAIN 480

QY 1416 PMDNSLYVLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAV 1475

Db 481 PMDNSYVLDNNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGKHAVQTTLESATAIAV 539  
QY 1476 SHNGVLYIAETDEKINRIRQVTTSGEISLVAGAPSGCDCNDANCDCFSGDDGYAKDAK 1535  
Db 540 SYSGVLYITETDEKINRIRQVTTDGEISLVAGIPSECDCKNDANCDCYQSGDGYAKDAK 599  
QY 1536 LNTPSSLAVCAGBELYVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLEFDTTGKH 1595  
Db 600 LNAOSSLAASPDGTLYIADLGNIRIRAVSKNKPLLNSMNFYEVASPTDQELYIFDINGTH 659  
QY 1596 LYTQSLPTGDYLYNFTYTGDDITLITDNNGNMNVNRDSTGMFLWLVPDQGVYWTMG 1655  
Db 660 QYTVSLVTGDYLYNFSYNDNDVTAVTDSNGNTLRIRRDPMFVRVSPDNQVWLTIG 719  
QY 1656 TNSALKSVTTQCHELAMMTYHGNSGLLATKSNENGWTTFFEYDSFGRLTNVTFPTGQVSS 1715  
Db 720 TNGCLKSMTAQGLELVLFYTHGNSGLLATKSDEGTWTTFFDYDSEGRLLTNVTFPTGVVTN 779  
QY 1716 FRSDTSSVHVQVETSSK-DDVTITNLSASGAFYTLLOQDVNRNSYYIGADGSLRLLAN 1774  
Db 780 LHGDMDKAITVDIESSSREEDVSITSNLSIDSFYTMVQDQLRNSYQIGYDGSRLIFYAS 839  
QY 1775 GMEVALQTEPHLLAGTVNPTVGNKRVNTPIDNGLNLVWRQKEARGQVTVFGRRRLRVH 1834  
Db 840 GLDSHYQTEPHVLAGTANPTVAKRNMTLPGENGNLVEWFRKEBQAQGVNVPGRKLRVN 899  
QY 1835 NRNLLSDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSSSRLNGVNVITYSPGGVIA 1894  
Db 900 GRNLLSVDFDRTTKTEKIYDDHRKFLRLIAYDTSGHPTLWLPSSKLMVNVITYSSSTGQIA 959  
QY 1895 GIORGIMSERMEYDQAGRITSRIEADGKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLS 1954  
Db 960 SIQRTTSEKVDYDSQGRIVSRVFADGKTWSYTYLEKSMVLLHLSQRQYIFEYDMWDRLS 1019  
QY 1955 SVTMPNVARQTLTETIRSVGYRNIYQPEGNASVIOQDFTEDGHLLHTFYLGTRRRVIYKY 2014  
Db 1020 AITMPSVARHTMTQIRSIGYRNIYNPPEPNASIIIDYNEEGLLQTAFLGTSRRVLFKY 1079  
QY 2015 GKLSKLAETLYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRETEB 2074  
Db 1080 RRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYQIGPLIDRQIFRSEB 1139  
QY 2075 GMVNARFDYNDNSFRVTSMAQVINETPLIDLYRYDDVSGKTEQFGKFGVIYYDINQII 2134  
Db 1140 GMVNARFDYSYDNSFRVTSMQGVINETPLIDLYQFDDISGKVEQFGKFGVIYYDINQII 1199  
QY 2135 TTAVMTHTKHEDAYGRMKEVQYEIFRSLMYMWTVOYDNMGRVVKKELKVGPIYANTTRYSY 2194  
Db 1200 STAVMTYTKHEDAHGRIKEIQYEIFRSLMYWITIOYDNMGRVTKREIKIGPFANITKYAY 1259  
QY 2195 EYDADGLOQTVSINDKPLWRYSYDLNGLHLLSPGNSARLTPLRYDIRDRITRLGDVQYK 2254  
Db 1260 EYDVGLOQTVVYLNKIMWRVNYDLNGLHLLNPSSSARLTPLRYDLRDRITRLGDVQYR 1319  
QY 2255 MDEDFLRQGGDIFEYNSAGLLIKAYNRAGSWVRVRYDGLGRRVSSKSSHHLQFFY 2314  
Db 1320 LDEDFLRQGRTEIFEYSSKGLLTRVYSGSGWTVIYRYDGLGRRVSSKTSLGHLQFFY 1379  
QY 2315 ADLTNPTKVTHLYNHSSSEITSLYYDLQGHLLFAMELSSGDEFYIACDNIGTPLAVFSCTG 2374  
Db 1380 ADLTPTTRITHVYNHSSSEITSLYYDLQGHLLFAMEISSGDEFYIASDNTGTPLAVFSSNG 1439  
QY 2375 LMIKQILYTAQEIMYMDTNPNFQIICGYHGGLYDPLTKLVNMGRRDYDLVLAGRWTSPDHE 2434  
Db 1440 LMLKQIQYTAQEIMYFDSNVDFQLVGFHGGLYDPLTKLHFGERDYDILAGRWTTPDIE 1499  
QY 2435 LWKHLSSSNVMPFNLYMFKNPNPISNSQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDM 2494  
Db 1500 IWKRI-GKDPAPFNLYMFRNNPNPASKIHDVKDYITDVNSWLVTFGFHLHNAIPGFVPKPF 1558  
QY 2495 DAMEPSYELIHTQMKTOEWDNKSILGVQCEVQKQKAFVTLERFDQLYGSTITSQOAP 2554  
Db 1559 DLTEPSYELV-----KSQQWEDVPIFGVQQQVARQAKAFLSLGKMAEV-----QVS 1605

QY 2555 KTKK-----PASSGSVFGKGVKFAKDKGRVTTDIISVANEDGRRVAAAILNHAHYLEN 2606  
Db 1606 RRKAGAEQSWLWFATVKSLLIGKGVMLAVSQGRVQTNVLNIANEDCIKVAAVLNNAFYLEN 1665  
QY 2607 LHFTIDGVDTHYFVKPGPSEGDLAJLGLSGRRRTLENGVNVTVSQINTVLNGRTRRYTDI 2666  
Db 1666 LHFTIEGKDTHYFIKTTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVVNGRTRRFADV 1725  
QY 2667 QLQYGALCLNTRYGTTLDDEKARVLELARQAVRQAWAREQQRLREGEGLRAWTEGEKQ 2726  
Db 1726 EMQFGALALHVRVYGMTLDEEKARILEQARQALARAWAREQQRVRDGEEGARLWTEGEKR 1785  
QY 2727 QVLSTGRVQYDGGFFVISVEQYPELSDSANNIHFMRQSEMGR 2769  
Db 1786 QLLSAGKVQYDGYVLSVEQYPELADSAANNIQFLRQSEIGKR 1828

Search completed: June 24, 2004, 16:21:35  
Job time : 100 secs



```
QY 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD 120
Db 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD 120

QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTTP 180
Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTTP 180

QY 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHAQENWL 240
Db 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHAQENWL 240

QY 241 LNSNIPLETRNLGKQPFGLTQDNLNLIEMDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS 300
Db 241 LNSNIPLETRNLGKQPFGLTQDNLNLIEMDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS 300

QY 301 PGYPLTSSVTYSPPPRPLPRSTFARPAFNKKPSKYNWKAALSAIVISATLVILLAYF 360
Db 301 PGYPLTSSVTYSPPPRPLPRSTFARPAFNKKPSKYNWKAALSAIVISATLVILLAYF 360

QY 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTEGK 420
Db 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTEGK 420

QY 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIY 480
Db 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIY 480

QY 481 RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDSGIWH 540
Db 481 RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDSGIWH 540

QY 541 LAFYNDGKSEVVSFLTTAI---ESVDNCPNSCYNGD-----CISGTCHCFLGFL---- 588
Db 541 LAFYNDGKSEVVSFLTTAI---ESVDNCPNSCYNGD-----CISGTCHCFLGFL---- 588

QY 589 GPDCCGRASCP-----VLCSGNGQYMKGRCLCHSGWKGAECVPTNQC 630
Db 589 GPDCCGRASCP-----VLCSGNGQYMKGRCLCHSGWKGAECVPTNQC 630

QY 601 GNELNRKNKPNQIKQKPKGKSELINRKASLADSGEYM-----CKVISKLG 646
Db 601 GNELNRKNKPNQIKQKPKGKSELINRKASLADSGEYM-----CKVISKLG 646

QY 631 IDVACSNHGTCTGTICINPGYKGESCEVDCMDPTCSGRGVCVRGECCHFCVWGWTNCE 690
Db 631 IDVACSNHGTCTGTICINPGYKGESCEVDCMDPTCSGRGVCVRGECCHFCVWGWTNCE 690

QY 647 NDSASAN---ITIV-----ESNEIITGMPASTEGAYVSSSPIRISVSTEGANTS 693
Db 647 NDSASAN---ITIV-----ESNEIITGMPASTEGAYVSSSPIRISVSTEGANTS 693

QY 691 TPRAT-----CLDQ-----CSGHGTFL-----PDTGLCSCDPSWTHGDCSIEIC 729
Db 691 TPRAT-----CLDQ-----CSGHGTFL-----PDTGLCSCDPSWTHGDCSIEIC 729

QY 694 SSTSTSTGTSHLVKCAEKEKTFVNGGBCFWVKDLNPSRYLCKCPNEFTGDRQCQNYVM 753
Db 694 SSTSTSTGTSHLVKCAEKEKTFVNGGBCFWVKDLNPSRYLCKCPNEFTGDRQCQNYVM 753

QY 730 AA 731
Db 754 AS 755
```

RESULT 2  
US-09-514-573-2  
; Sequence 2, Application US/09514573  
; Patent No. 6500941  
; GENERAL INFORMATION:  
; APPLICANT: Schaefer, Gabriele M.  
; APPLICANT: Sliwowski, Mark  
; TITLE OF INVENTION: Gamma-Heregulin  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-09-514-573-2
```

```
Query Match 20.1%; Score 2989; DB 4; Length 768;
Best Local Similarity 76.5%; Pred. No. 1.2e-208;
Matches 598; Conservative 22; Mismatches 84; Indels 78; Gaps 10;

QY 1 MDVKERKPYRSILTRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60
Db 1 MDVKERKPYRSILTRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60

QY 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD 120
Db 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD 120

QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTTP 180
Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTTP 180

QY 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHAQENWL 240
Db 181 PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHAQENWL 240

QY 241 LNSNIPLETRNLGKQPFGLTQDNLNLIEMDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS 300
Db 241 LNSNIPLETRNLGKQPFGLTQDNLNLIEMDILGASRHDGAYS DGHFLFKPGGTSPLFCTTS 300

QY 301 PGYPLTSSVTYSPPPRPLPRSTFARPAFNKKPSKYNWKAALSAIVISATLVILLAYF 360
Db 301 PGYPLTSSVTYSPPPRPLPRSTFARPAFNKKPSKYNWKAALSAIVISATLVILLAYF 360

QY 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTEGK 420
Db 361 VAMHLFGLNWLQPMEGQMEYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTEGK 420

QY 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIY 480
Db 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIY 480

QY 481 RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDSGIWH 540
Db 481 RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPSSSHETGFIQYLDSGIWH 540

QY 541 LAFYNDGKSEVVSFLTTAI---ESVDNCPNSCYNGD-----CISGTCHCFLGFL---- 588
Db 541 LAFYNDGKSEVVSFLTTAI---ESVDNCPNSCYNGD-----CISGTCHCFLGFL---- 588

QY 589 GPDCCGRASCP-----VLCSGNGQYMKGRCLCHSGWKGAECVPTNQC 630
Db 589 GPDCCGRASCP-----VLCSGNGQYMKGRCLCHSGWKGAECVPTNQC 630

QY 601 GNELNRKNKPNQIKQKPKGKSELINRKASLADSGEYM-----CKVISKLG 646
Db 601 GNELNRKNKPNQIKQKPKGKSELINRKASLADSGEYM-----CKVISKLG 646
```

QY	631	IDVACSNHGTCTGTCTICNPGYKGESCEEVD	CMDPTCSGRGYCVRGEC	HCFVGWGGTNC	E 690
Dd	647	NDSASAN---	ITIV-----ESNEIITGMPASTE	GAYVSSESPIRISVSTEGANTS	E 693
QY	691	TPRAT-----CLDQ-----	CSGHGTFL----	PDTGLCSCDP	SWTGHDCSIHC E 729
Dd	694	SSTSTSTGTSHLVKAEKEKTFCVNGGECF	MVKDLNSPBYLCKCPNEFTGDRCQNYVM		E 753
QY	730	AA 731			
Dd	754	AS 755			

```

RESULT 3
US-08-891-845-4
; Sequence 4, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb fl
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-891-845-4

```

Query Match 20.0%; Score 2982; DB 3; Length 560;  
Best Local Similarity 99.8%; Pred. No. 2.2e-208;  
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSETLKAYDQDARLAYGSRV	60
Dd	1	MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSETLKAYDQDARLAYGSRV	60
QY	61	KDIVPQAEFFCRTGANFTLRELGLBEVTPPHGTLYRTDIGLPQCYSMGAGSDADMEAD	120
Dd	61	KDIVPQAEFFCRTGANFTLRELGLBEVTPPHGTLYRTDIGLPQCYSMGAGSDADMEAD	120
QY	121	TVLSPHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQN HARLTRPP	180
Dd	121	TVLSPHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQN HARLTRPP	180

QY	181	PPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHQENWL	240
Db	181		
QY	241	LNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPFLFCTTS	300
Db	241		
QY	301	PGYPLTSSVTYSPPPRPLPRSTFARPAFNLKKPSKYNCKWKAALSAIVISATLVILLAYF	360
Db	301		
QY	361	VAMHLFGLNWLQPMEGQWYEITEDTASSWPVPTDVSYPGGTGLETDRKGKGTTEGK	420
Db	361		
QY	421	PSSFFPEDSFIDSGEIDVGRASQKIPPGTFWRQSVFIDHPVHLKFNYSLGKAALVGIYG	480
Db	421		
QY	481	RKGLPPSHTQDFVELLDGRRLLTQEARSLGTPRQSRGTVPPSSHETCFIQYLDSGIWH	540
Db	481		
QY	541	LAFYNDGKESEVVSFLTITAI	560
Db	541		

RESULT 4  
 US-09-514-573-4  
 ; Sequence 4, Application US/09514573  
 ; Patent No. 6500941  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schaefer, Gabriele M.  
 ; APPLICANT: Sliwowski, Mark  
 ; TITLE OF INVENTION: Gamma-Hersgulin  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 460 Point San Bruno Blvd  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Winpatin (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/514,573  
 ; FILING DATE: 28 FEB 2000  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/891845  
 ; FILING DATE: 10 JULY 1997  
 ; APPLICATION NUMBER: 60/021640  
 ; FILING DATE: 07/12/96  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Lee, Wendy M.  
 ; REGISTRATION NUMBER: 40,378  
 ; REFERENCE/DOCKET NUMBER: P1043  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415/225-1994  
 ; TELEFAX: 415/952-9881  
 ; TELEX: 910/371-7168  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 560 amino acids  
 ; TYPE: Amino Acid  
 ; TOPOLOGY: Linear  
 ; US-09-514-573-4

Query Match 20.0%; Score 2982; DB 4; Length 560;  
Best Local Similarity 99.8%; Pred. No. 2.2e-208;  
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVKERKPYRSLRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60  
DB 1 MDVKERKPYRSLRRRDAERRYTSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60

QY 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDIGLPQCGYMGAGSDADMEAD 120  
DB 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLTYRTDIGLPQCGYMGAGSDADMEAD 120

QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEHENTETDHPGGLQNHARLRTPP 180  
DB 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEHENTETDHPGGLQNHARLRTPP 180

QY 181 PPLSHAHTPNQHHAAASINSLNRGNFTFRSNPSPAPTDHLSLSEPPAGGAQEPAAHAQENWL 240  
DB 181 PPLSHAHTPNQHHAAASINSLNRGNFTFRSNPSPAPTDHLSLSEPPAGGAQEPAAHAQENWL 240

QY 241 LNSNIPLETRNLGKQPFGLTQDNLIEDILGASRHDGAYSDDHFLFKPGGTSPLFCTTS 300  
DB 241 LNSNIPLETRNLGKQPFGLTQDNLIEDILGASRHDGAYSDDHFLFKPGGTSPLFCTTS 300

QY 301 PGYPLTSSVSPPPRPLPRSTFARPAFNLKPKSKYCNWKCAALSIAVISATLVILLAYF 360  
DB 301 PGYPLTSSVSPPPRPLPRSTFARPAFNLKPKSKYCNWKCAALSIAVISATLVILLAYF 360

QY 361 VAMHLFGLNWLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLTDPDRKGKGTTEGK 420  
DB 361 VAMHLFGLNWLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLTDPDRKGKGTTEGK 420

QY 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAALVGIY 480  
DB 421 PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAALVGIY 480

QY 481 RKGLPPSHTQDFVELLDGRRLLTQARSLEGTPRQSRGTVPSPSSHETGFIQYLDSDGIWH 540  
DB 481 RKGLPPSHTQDFVELLDGRRLLTQARSLEGTPRQSRGTVPSPSSHETGFIQYLDSDGIWH 540

QY 541 LAFYNDGKESEVVSFLTTAI 560  
DB 541 LAFYNDGKESEVVSFLTTAI 560

RESULT 5

US-08-891-845-10

; Sequence 10, Application US/08891845

; Patent No. 6096873

; GENERAL INFORMATION:

; APPLICANT: Schaefer, Gabriele M.

; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Heregulin

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/891,845

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/021640

; FILING DATE: 07/12/96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P1043  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1994  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 501 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-08-891-845-10

Query Match 10.6%; Score 1575; DB 3; Length 501;  
Best Local Similarity 64.5%; Pred. No. 5.6e-106;  
Matches 332; Conservative 22; Mismatches 83; Indels 78; Gaps 10;

QY 268 MDILGASRHDGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSTVYSPPPRPLPRSTFARPA 327  
DB 1 MDILGASRHDGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSTVYSPPPRPLPRSTFARPA 60

QY 328 FNLKKPSKYCNWKCAALSIAVISATLVILLAYFVAMHLFGLNWLQPMEGQMYEITEDTA 387  
DB 61 FNLKKPSKYCNWKCAALSIAVISATLVILLAYFVAMHLFGLNWLQPMEGQMYEITEDTA 120

QY 388 SSWPVPTDVSLYPSGGTGLTDPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRRASQKIP 447  
DB 121 SSWPVPTDVSLYPSGGTGLTDPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRRASQKIP 180

QY 448 PGTFWRSQVFIIDHPVHLKFNVSGLKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEA 507  
DB 181 PGTFWRSQVFIIDHPVHLKFNVSGLKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEA 240

QY 508 RSLEGTPROSRGTVPSPSSHETGFIQYLDSDGIWHLAFYNDGKESEVVSFLTTAI ---ESVD 564  
DB 241 RSLEGTPROSRGTVPSPSSHETGFIQYLDSDGIWHLAFYNDGKESEVVSFLTTAI ALPRLK 300

QY 565 NCPNSNCYNGD ---CISGTCHCFLGFL ---GPDCCGRASCP --- 598

DB 301 EMKSQESAGSKLVLRCTSSSEYSSLRFKFWKNGNELNRKNKPNQIKQKPGKSELRLIN 360

QY 599 -VLCSGNGQYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTGTCTICNPGYKGESC 657  
DB 361 KASLADSGEYM ---CKVISKLGNDSASAN ---ITIV ---ESN 393

QY 658 EEVDCMDPTCSGRGVCVRGECHCFVGGWGTNCETPRAT ---CLDQ ---CSG 702  
DB 394 EIITGMPASTEGAYVSESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTFVCVNG 453

QY 703 HGTF ---PDTGLCSCDPSWTGHDCSIIEICAA 731  
DB 454 GECFMVKDLSNPSRYLCKCPNEFTGDRCONYVMAS 488

RESULT 6

US-09-514-573-10

; Sequence 10, Application US/09514573

; Patent No. 6500941

; GENERAL INFORMATION:

; APPLICANT: Schaefer, Gabriele M.

; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Heregulin

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
FILING DATE: 28 FEB 2000  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/891845  
FILING DATE: 10 JULY 1997  
APPLICATION NUMBER: 60/021640  
FILING DATE: 07/12/96  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P1043  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 501 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-09-514-573-10

Query Match 10.6%; Score 1575; DB 4; Length 501;  
Best Local Similarity 64.5%; Pred. No. 5.6e-106;  
Matches 332; Conservative 22; Mismatches 83; Indels 78; Gaps 10;

QY 268 MDILGASRHDGAYS DGHFLFKPGGTSPLFCTTSPGYPLTSSVTYSPPPRPLPRSTFARPA 327  
D 1 MDILGASRHDGAYS DGHFLFKPGGTSPLFCTTSPGYPLTSSVTYSPPPRPLPRSTFARPA 60  
QY 328 FNLKPKSKYCNWKCAALSAIVISATLVILLAYFVAMHLFGLNWHLQPMEGQMEYEITEDTA 387  
D 61 FNLKPKSKYCNWKCAALSAIVISATLVILLAYFVAMHLFGLNWHLQPMEGQMEYEITEDTA 120  
QY 388 SSWPVPTDVSLYPSGGIGLETDPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRRASQKIP 447  
D 121 SSWPVPTDVSLYPSGGIGLETDPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRRASQKIP 180  
QY 448 PGTFWRSQVFIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEA 507  
D 181 PGTFWRSQVFIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEA 240  
QY 508 RSLGTPRQSRGTVPSSSHETGFIQYLDSGIWHLAFYNDGKSEVVSFLTATAI---ESVD 564  
D 241 RSLGTPRQSRGTVPSSSHETGFIQYLDSGIWHLAFYNDGKSEVVSFLTATAIAPPLRK 300  
QY 565 NCPNSCYNGD-----CISGTCHCFLGFL-----GPDGGRASCP----- 598  
D 301 EMKSQESAAGSKLVLRCTSSSEYSSLRFKWFKNGNELNRKNKPQNIKIQKPKKSELIN 360  
QY 599 -VLCSGNGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTCITGTICINPGYKGESC 657  
D 361 KASLADSGEYM-----CKVISKLGNDSASAN---ITIV-----ESN 393  
QY 658 BEVDCMDPTCSGRGVCVRGECHCFVWGWTNCETPRAT-----CLDQ-----CSG 702  
D 394 EIITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTEFCVNG 453  
QY 703 HGTFEL-----PDTGLCSDPSPSWTGHDCSIEICAA 731  
D 454 GECFMVKDLNPSRYLCKPNEFTGDRCQNYVMAS 488

RESULT 7

US-09-796-575-2

; Sequence 2, Application US/09796575

Patent No. 632671  
GENERAL INFORMATION:  
APPLICANT: Genesegues, Inc.  
TITLE OF INVENTION: NANOCAPSULE ENCAPSULATION SYSTEM AND METHOD  
FILE REFERENCE: G332.12-0001  
CURRENT APPLICATION NUMBER: US/09/796,575  
CURRENT FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: US 60/185,282  
PRIOR FILING DATE: 2000-02-28  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 2200  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-796-575-2

Query Match 4.4%; Score 649.5; DB 4; Length 2200;  
Best Local Similarity 22.3%; Pred. No. 1.7e-37;  
Matches 310; Conservative 166; Mismatches 492; Indels 423; Gaps 64;

QY 545 NDGKESEVVSFLTITTAIESVDNCPNSCYNGDCISGTRCHCFLGLGPDGGRASCPVLCS-G 603  
D 208 DDGFTGEDCSQLA-----CPSDCNDQKCVNGVICIFEGYAGADCSREICPVPCSEE 259  
QY 604 NGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTCITGTICINPGYKGESCEEVDCM 663  
D 260 HGTCVDGLCVCHDGFAGDDCNKPL--CLN-NCYNRGRGVENECVCEGTGDCSELICP 316  
QY 664 DPTCSGRGVCVRGECHCFVWGWTNCETPRATCLDQCSHGHTFLPDTGLCSCDPSWTGHD 723  
D 317 N-DCFDRGRCTINGTCYCEEGETGEDCGKP--TCPHACTQGRC--EEGQCVCDEGFAGLD 371  
QY 724 CSIEICAADCGHGVCGVGTGTCRCEDEGMMGAACDQORACHPRCAEHGTCRDGKCECSPGWNG 783  
D 372 CSEKPCADCHNRGRCVDRGCECDDGFTGADCGELKCPNGCSGHGRCVNGQCVCDEGYTG 431  
QY 784 EHCT-----IAHYLDRVVKEG-----CPGLCNGNGRCITLDLNGWHCV 820  
D 432 EDCSQLPCPNDCHSRGRCV-EGKCVCEQGFGKYDCSDMSPCNDCHQHGRG---VNGM-CV 486  
QY 821 COLGWRGAGC-----DTSNETACGDSKDNDDGLV--DCMDPDC-----CL--Q 860  
D 487 CDDGYTGEDCDRQCPDCNRGLCVDGQCVCEDEGFTGPDCAELSCPNDCHGQGRCVNGQ 546  
QY 861 PLCH---INPLCLGSPNPLDIQETQVPVSQQNLHSHFVDRIKFLVGRDSTHIIPGENPFD 917  
D 547 CVCHGFMGKDCKEQRCPSPDCHGQGRCVDGQCICHEGF-----TGLDC----- 589  
QY 918 GHAC-----VIRGQVNTSDGTPLVGNISFVNPNLFGYTTISRQDGSFDLVTNGGI 968  
D 590 GQHSQPSDCNNLQCVSGRCICNEG--YSGEDCSEVSPK-----DLV----- 630  
QY 969 SIILRPERAPFITQ--EHTLWLPWD-RFFVMETIIMRHEENEIPSCDLSNFARPNPVVSP 1025  
D 631 -----VTEVTEETVNLAWDNEMRVTEYL-----VYTP 658  
QY 1026 SPLTSPASSCAEKGPVPEIQALQEEIISGCKMRLSYLSSRTPGYKSVLRISL---THP 1082  
D 659 T-----HEGLEMQFRVPGDQTS-----TIQELEPGVEYFIRVFAILENKK 700  
QY 1083 TIPFNLMKVHLMVAVEGRLFKRWFAAAPDLSSYFIWKTIDV-----YNQKVFG- 1130  
D 701 SIPVSARVATYLPAPGLKFK---SIKETSVEVEWDPLDIAFETWEIIFRNMNKEDEGE 756  
QY 1131 -----LSEAFVSVGYEYESCPDLI-----LWEKRTTVLQG---YEI-----D 1164  
D 757 ITKSLRPPETSYRQTGLAPGQVEYISLHIVKNTRGPGVKRVTTRLDAPSQIEVKDVTD 816  
QY 1165 ASKLGW--SLDKHHALNIQSGI-----LHKNGENQFVSQQPPVIGSIMGNRRR 1213  
D 817 TTALITWFKPLAIEDGIELTYGIKDVPGDRRTTIDLTEDENQY-----SIGNLKPD-TEY 869



QY 1564 RKNKPFNTQNMVELSSPIDQELYLFTDTTGKHLTYQSLPTGDYLYNFYVTGDDITLITD 1623  
Db 1176 GWDALKLN-----WTAPEGAYEYFFIQVEADTVEAAQ 1208  
QY 1624 NNGNMVNVRRDSTGMPLWLVPDGOVYVWMTGNSALK-SVTTQGHSLAMMTYHGNSGLL 1682  
Db 1209 N-----LTVPGGLRSTDLPLGLKAATHYTITIRGVTQDFSTPLSVEVL 1251  
QY 1683 ATKSNENGWTTFEYDPSFRLTNVTFPTQVSSFRSDTSSVHVQVETSSKDDVTITNL 1742  
Db 1252 TEEVPDMGNLTVEVSWDALRLNWTTPDGTQDFT-----IQVQEA----- 1292  
QY 1743 SASGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGT 1790  
Db 1293 -----DQVEEAHNLTVPGSLR-----SMEI-----PGLRAGT 1319  
RESULT 9  
PCT-US95-11684-2  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE  
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING  
; TITLE OF INVENTION: AND USING SAME  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: The Scripps Research Institute, Office of  
; ADDRESSEE: Patent Counsel  
; STREET: 10666 North Torrey Pines Road, TPC 8  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 14-SEP-1995  
; APPLICATION NUMBER: PCT/US95/11684  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/308,359  
; FILING DATE: 16-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Logan, April C.  
; REGISTRATION NUMBER: 33,950  
; REFERENCE/DOCKET NUMBER: BEC0019P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-554-2937  
; TELEFAX: 619-554-6312  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2199 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-11684-2  
Query Match 4.4%; Score 648.5; DB 5; Length 2199;  
Best Local Similarity 21.4%; Pred. No. 2e-37;  
Matches 280; Conservative 169; Mismatches 474; Indels 385; Gaps 55;  
QY 566 CPSNCYGNGDCISGTCFLGLPDCGRASCPVLCSGNQYMKGRCLCHSGWKGAECD- 624  
Db 314 CPNDCFDRGRCINGTCYCEBGTGEDCGKPTCPHACTQGRCEEGQVCDEGFAVDCE 373  
QY 625 --VPTN-----QCID-----VACSNHGTCITGTCTCNCNPGYKGESC 657  
Db 374 KRCPADCHNRGRVCVDRCEDDGFTGADCGELKCPNGCSGHGRVNGQVCVDEGYTGDC 433

QY 658 EEVDCMDPTCSGRGVCVRGECHCFVWGWTNCETPRATCLDQCShGHTFLPDITGLCSDDP 717  
Db 434 SQLRCPN-DCHSRGRVCVEGKVCCEQGFKGYDCSD--MSCPNDCHQHGRCV--NGMCVCDD 488  
QY 718 SWTGHDCSIEICAADCGGHGVCVGTGTCRCEGMMGAACDQRAHPRCAEHGTCRDKCEC 777  
Db 489 GYTGEDCRDQCPRDCSNRGLCVGQVCCEGFTGPDCAELSCPNDCHGQRCVNGQCVC 548  
QY 778 SPWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHVCVQLGWRGAGCDTSMETA 837  
Db 549 HEGFMGDKC-----EQRCPSDCHGQGR-----VDG-QCICHEGFTGLDCG--QHS 592  
QY 838 CGDSKNDGDGLVDCMDPDCLQPLCHINPLCLG-----SPNPLDIIQETQVPVSQNL 891  
Db 593 CPS-----DCNNLGQCVSGRCICNEGYSGEDCSEVSP-PKDLV--VTEVTEETV 638  
QY 892 HSFYDR-----IKFLVGRDSTHIIPGENPFDDGGHACVIRGQVMTSDGTPLV-----GVN- 940  
Db 639 NLAWDNEMRVTEYLVVYTPH-----EGGLEMOFR--VPGDQTSTIIRELEPGVEY 687  
QY 941 ----ISFVNNPLFGYTIISRDGSDFLVTNGGISIILFERAPERFITQEHTLWLPWDRFFVM 996  
Db 688 FIRVFAILENKK-SIPVSARVATYLPAPG-----LKFPSIK-ETSVEVEWDPLDIAFET 740  
QY 997 ETIMRHEENEIPSCDLSNFARPNPVSPSLTSFASCAEKGPVPEIQALQEBISISG 1056  
Db 741 WEIIFRNMNKEDEGEITKSLRPE-----TSYRQTGLAPG-----QEYELS- 781  
QY 1057 CKMRLSYLSSRTPGYKSVLRISLTHPTIPFNLKMKVHLMVAVEGRLFRKWFAPADLSYF 1116  
Db 782 --LHIVKNNTRGPLKRVTTTRELDAPS-----QIEVKDVTDTTALITWFKPLAED--- 830  
QY 1117 IWDXTDVYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKH 1176  
Db 831 -----GIELTYGIKDVPG-----DRTIDLTEDENQYSIGNLKPDTE 867  
QY 1177 HALNIQSGILHKGNGENQFVSQQPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPALT 1236  
Db 868 YEVSIL---ISRRGD-----MSSNP-----AKETFT 889  
QY 1237 CGSDGSLYVGDFNYIRRIFFPSGNVTNILELRNKDFRSHSPAHKYLATDPMGAVPLSD 1296  
Db 890 TGLDAP-----RNLRRVSQDINSITL-----EWRNGKAAIDSYRIKYAPISGG----- 932  
QY 1297 SNRRVFKIKSTVAVKLVKNSEVAVGTGDCQLPDDTRCGDGGKATEATLTNPRGITVD 1356  
Db 933 -----DHAEDVVPKSQLA-----TTKTTTLTGLRPGT-- 958  
QY 1357 KFGLIYFVDGTMIRRIDQNGIISTLLGSNDLTARPLSCDSVMDISQVRLEWPTDLA-- 1413  
Db 959 EYGI-----GVSARKEDKESNPATINAATELDTPKDLQVSETAETS-LTLLWKTPLAKFD 1012  
QY 1414 --IN-----PMDNSLYVLDNNVVLQISENHQVRIVAGRPMHC----- 1448  
Db 1013 RYRLNYSPLPTGQWVGVLPRNTTSYVLRG--LEPGQEVNVLTAEKGRHKSKPARVKAS 1069  
QY 1449 --QVPGIDHFLLSKVAIHATLESATAL--AVSHNGVLYIAETDE-KKINRIRQVTTSGEI 1503  
Db 1070 TEQAPLENLTVTEVGWDGLRLNWTAAOQAYEH---FIQVQEAANKVEAARNLTVPF-- 1123  
QY 1504 SLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLNTPSLAVCADGELYVADLGNIRIRFI 1563  
Db 1124 SLRAVDIPGLKAATPYTVTSYVIGQY-----RTPVLSABEASTGE--TPNLGEVVVAEV 1175  
QY 1564 RKNKPFNTQNMVELSSPIDQELYLFTDTTGKHLTYQSLPTGDYLYNFYVTGDDITLITD 1623  
Db 1176 GWDALKLN-----WTAPEGAYEYFFIQVEADTVEAAQ 1208  
QY 1624 NNGNMVNVRRDSTGMPLWLVPDGOVYVWMTGNSALK-SVTTQGHSLAMMTYHGNSGLL 1682  
Db 1209 N-----LTVPGGLRSTDLPLGLKAATHYTITIRGVTQDFSTPLSVEVL 1251  
QY 1683 ATKSNENGWTTFEYDPSFRLTNVTFPTQVSSFRSDTSSVHVQVETSSKDDVTITNL 1742

Db 1252 TEEVPMGNLTTEVSWDALRLNWTTPDGYDQFT-----IQVEA----- 1292

QY 1743 SASGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGT 1790

Db 1293 -----DOVEEAHNLTVPGSLR-----SMEI-----PGLRAGT 1319

RESULT 10

US-08-793-273C-4

; Sequence 4, Application US/08793273C

; Patent No. 6482410

; GENERAL INFORMATION:

; APPLICANT: Crossin, Kathryn L.

; APPLICANT: Phillips, Greg

; APPLICANT: Prieto, Anne L.

; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE ATTACHMENT AND

; TITLE OF INVENTION: NEURITE OUTGROWTH, AND METHODS OF MAKING SAME

; FILE REFERENCE: BEC00225

; CURRENT APPLICATION NUMBER: US/08/793,273C

; CURRENT FILING DATE: 1997-05-22

; PRIOR APPLICATION NUMBER: PCT/US95/11684

; PRIOR FILING DATE: 1995-09-14

; PRIOR APPLICATION NUMBER: 08/308,359

; PRIOR FILING DATE: 1994-09-16

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Patentin ver. 2.1

; SEQ ID NO 4

; LENGTH: 1810

; TYPE: PRT

; ORGANISM: Gallus gallus

US-08-793-273C-4

Query Match 4.3%; Score 641; DB 4; Length 1810;

Best Local Similarity 20.4%; Pred. No. 5.le-37;

Matches 333; Conservative 219; Mismatches 563; Indels 520; Gaps 72;

QY 543 FYN DGKESVVSFLTITAESVDNCPNCGNGDCISGTCHEFLGFLGPDCCGRASCPVLCS 602

Db 332 FCEEGYTGDCGELT-----CPNNGNCGRCENGCLVCHEGFVGGDCSQKRCPTCN 383

QY 603 GNGQYMKGRCLCHSGWKGAECD---VPTNQCIDVACNSHGTCTGTCTICNPGYKGCSEEE 659

Db 384 NRGRVCDGRVCHEGYLGEDCGELRCND-----CHNRGRINGQCVCDGEGFICGCE 437

QY 660 VDCMDPTCSGRVCVRGECHCFVWGWTNCETPRATCLDQCSGHGTFPLPDTGLCSCDPSW 719

Db 438 LRCPN-DCQQRGRINGQCECHGFCGEDCGELR--CPNDCNSHGRCV--NGQCVCDEGY 492

QY 720 TGHDCSIEICAADCGHGVCGVGTCTCEDGWMGAACDQACHPRCAEHGTCRDGKCECSP 779

Db 493 TGEDCGELRCNDCHNRGRVCGRCVCGCDGFMGDCGELSCPNDCCHQGRVCDGRVCHE 552

QY 780 GWNGEHCCTIAHYLDRVVKEGCPCGLCNGRCRCTLDLNGWHVCQLGWRGAGCDTSMETACG 839

Db 553 GFTGEDCR-----ERSCPNDCNNVGR-----VEG-RCVCEEYMG----- 587

QY 840 DSKDNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQFTQVPVSQQNLHSFYDRI- 898

Db 588 -----IDCS-----VSP-----PTGLTVTNVTDKTV---NLEWKHENLV 619

QY 899 -KFLV-----GRDSTHIIPG-----NPFDDGGHACVIRGQVMTSDGTPLVGVNISF 943

Db 620 NEYLVTVYPTSSGGLDLQFTVPGNQTSATIHLEPGVEYFIRVFAILKNKKS----- 672

QY 944 VNNPLFGYTTISRQDGSFDLVNNGGISILRFERAPFITQEHTLWLPWDRF---FVMEI 1000

Db 673 -----PVSARVATYLPAGE-----LKFS-----VRETSQVQVEMDPLSISFDGWELV 715

QY 1001 MRHEENEIPSCDL-SNFARPNPVSRPLTSFASSCAEKGPVPEIQALQEEISISGCKM 1059

Db 716 FRNMQKDDNGDITSSLRPE-----TSYMQPGLAPG-----QQYNVS---L 754

QY 1060 RLSYSSRTPGYKSVLRISLTHPTTTPFNLMKVHLMVAVEGRLEFRKWFAAAPDLSYFYFWD 1119

Db 755 HIVKNTRGPGLSRVITTKLDAPS-----QIEAK-----D 784

QY 1120 KTDVYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTTVLQGYE-----IDA 1165

Db 785 VTDT-----TALITWSKPLAEIEGIELTYGPKDVPDGRDRTIDL 822

QY 1166 SKLGGWSLDKXHALNIQSGILHKNGENQFVSQQPPVIGSIMNGRRRS-----ISCPSC 1220

Db 823 SE-----DENQY-----SIGNLRPHTEYEVTLISR 849

QY 1221 NGLADGNKLLAPVALTCGSDGSLYVGFNYIRRIFFPSGNVTNILELRNKDFRHSHPAHK 1280

Db 850 RGDMESDPA-----KEVFVTDLDAPRNLKRVSQTDNSITL---EWKFSHANIDN 895

QY 1281 YYLATDPMGAVFLSDSNSRRRVFKIKSTVVVKLVKNSEVVAGTGDQCLPDDTRCGDGG 1340

Db 896 YRIKFAPISGG-----DHTELTVPKGNQA----- 919

QY 1341 KATEATLTNPERGITVDKFLIYFVDTMIRRIDQNGIISTJLGSNDLTSARPLSCDSVMD 1400

Db 920 -TTRATLTGLRP-----GTEYGIGVTAVRQ-DRESAPATINAGTDLNPKDLEVSDDPTE 971

QY 1401 ISQVRLEWPTDLA-----INPMDNSLYVDNNVVLQISENHQVRIV 1441

Db 972 -TTLSLRWRPVAKFDRYRLTYVSPSGKKNEMEIPVDSTSFILRG---LDAGTEYTTISLV 1027

QY 1442 AGRPMHCQVP-----GIDHFLLSKVIAIHATLESATALAVSHNGV 1480

Db 1028 AEKGRHKSKPTTIKGSTEEEPBELGNLSVSETGWDGFQLTWTAADGAYE-----NFV 1078

QY 1481 LYIAETDEKKINRIRQVTTSGEISLVAGAPS---GDCCKNDANDCDFSGDDGYAKDAKL 1536

Db 1079 IQVQQSDNPE-----ETWNIIVPGGQHSVNTGLKANTPYNVTLYGVIRGYRTKPLY 1130

QY 1537 NTPSSLAVCADGELYVADL-----GNIRIRFIRKNK---PF-----L 1570

Db 1131 VETTGAHPEVGELTVSDITPESFNLSWTTTNGDFDAFTIEIDSNRLLPEMFEFISGNS 1190

QY 1571 NTQNMVELSPIDQELYLFDTTGKHLV-----TQSLPTGDYLYNFYTYTGDGDTL 1620

Db 1191 RTAHISGLSPSTDFIVLYGIS--HGFRTOAISAAATTEAEPEVDNLLVSDATPDGFRLT 1248

QY 1621 ITDNNG-----NMVNVRDSTGMPLWLVP-----DGQVYVWVTMG 1656

Db 1249 WTADDGVFDSFVLKIRDTKRSD--PLELIVPGHERTHDITGLKEGTEYEIELYGVSSGR 1306

QY 1657 NS-ALKSVTTQCHELAMTYHGN-SGLLATKSNEN---GWTTFEYDSFGRLTNVTPPT 1710

Db 1307 RSQPINSVAT-----TVVGSPPKISFSDITENSARVSWTP-----PR 1343

QY 1711 QOVSSFRSDTSSVHVQVETSSKDDVTITNLSASGAFYTLQDQVRNSYYIGADGSLRL 1770

Db 1344 SRVDSYRVS-----YVPIITGTPNVTVVDGSKTR-----TKLVKLVP-----GVDYVNI 1388

QY 1771 LLANGMEVALQTEPHLLAGTVNPTVGRN--VTLPIDNGLNLEWVR-----QRK 1817

Db 1389 ISVKGFE---ESEP--ISGILKTALDSPSGLVVMNITDSEALATWQPAIAAVDNYIVSYS 1443

QY 1818 EQARGQVT--VFGRRLRVHNRNLLSLDFDRVTRTEKIYDHRKFTLRLILDQA--GRPSL 1873

Db 1444 SEDEPEVTQMVSGNTVE-YDLNGLRPATEYTLRVHAVKDAKSETLSTQFTTGLDAPKDL 1502

QY 1874 WSPSSRLNGVNVITYSPGGYIAGIQGIMSERMEYDQAGRITSIFADGKTWSYTYLEKSM 1933

Db 1503 SATEVQSETAVITWRPP-----RAPVTDYLLTYESIDGRVKEVILDPETTSYTLTE--- 1553

QY 1934 VLLLSQRQYIFEFDKNDRLSSVTMENVARQTLTIRSVGYRNIYQPPPEGNASVIOQDFT 1993

Db 1554 ---LSPSTQYTVKLQALSR---SMRSKMIQTFTTTGLLY----PYPKDCSQALLNGEV 1602

QY 1994 EDGHLHHTFYLGTR 2008

Db 1603 TSG--LYTIYLANGDR 1615

RESULT 11

PCT-US95-11684-4

Sequence 4, Application PC/TUS9511684

GENERAL INFORMATION:

APPLICANT: THE SCRIPPS RESEARCH INSTITUTE

TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE

TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING

TITLE OF INVENTION: AND USING SAME

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: The Scripps Research Institute, Office of

ADDRESSEE: Patent Counsel

STREET: 10666 North Torrey Pines Road, TPC 8

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/11684

FILING DATE: 14-SEP-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/308,359

FILING DATE: 16-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Logan, April C.

REGISTRATION NUMBER: 33,950

REFERENCE/DOCKET NUMBER: BEC0019P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-554-2937

TELEFAX: 619-554-6312

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1810 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-11684-4

Query Match 4.3%; Score 641; DB 5; Length 1810;

Best Local Similarity 20.4%; Pred. No. 5.1e-37;

Matches 333; Conservative 219; Mismatches 563; Indels 520; Gaps 72;

QY 543 FYNDGKESEVWFLTTAIESVDNCPNSCYNGDCISGTCCHFLGFLGPCDGRASCPVLCS 602

Db 332 FCEEGYTGEDCGELT-----CPNNCNGNRCENGCLVCVCHGCVGDCSOKRCPKTCN 383

QY 603 GNGQYMKGRCLCHSGWKGAEC--VPTNQCIDVACSNHGTCITGTICNPGYKGESCEE 659

Db 384 NGRGCVDRGCVCHGVLGEDCGELRCPND-----CHNRGRGCVGCVGDEFGEDCGE 437

QY 660 VDCMDPTCSGRGVCVRGECHQFVGWGGTNCETPRATCLDQCSGHGTFPLPDTGLCSDPSW 719

Db 438 LRCPN-DCQQRGRCINGQCEHGFIGEDCGELR--CPNDCNSHGRCV--NGQCVDEGY 492

QY 720 TGHDCSIEICAADCGHGVCGGTCTCRCECDGWMGAACDQACHPRCAEHGTCRDGKCECSP 779

Db 493 TGEDCGELRCPNDCHNRGRCVGRGCVCDNGFMGEDCGELSCPNDCHQHGRGCVDRGCVCHE 552

QY 780 GWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCRTLDLNGHCVCOLGWRGAGCDTSMETACG 839

Db 553 GFTGEDCR-----ERSCPDCNNVGRG---VEG-RCVCEEYMG----- 587

QY 840 DSKDNDGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVPSQQNLHSFYDRI- 898

Db 588 -----IDCSD-----VSP-----PTGLTVTNVTDKTV-----NLEWKHENLV 619

QY 899 -KFLV-----GRDSTHIIPGE-----NPFDCGHACVIRGOVMTSDGTPLVGVNISF 943

Db 620 NEYLVTYVPTSSGGLDLQFTVPGNQTSAIHELEPGVEYFIRVFAILKNKKS1----- 672

QY 944 VNNPLFGYTIISRDGSDFLVNTNGGISIIILRFRAPFITQEHTLWLPWDRF---FVMEITII 1000

Db 673 -----PV SARVATYLPAGE-----LKFKS-----VRETSVQVEWDPLS1SFDGWELV 715

QY 1001 MRHEENEIPSCDL-SNFARPNPVVSPSPLTSFASCAEKGP1VPEIQALQEE1SISGCKM 1059

Db 716 FRNMQKKDDNGDITSSLKRP-----TSYMQPLAPG-----QQYNVS---L 754

QY 1060 RLSYLSSRTPGYKSVLRISLTHPTIPFNLKMKVHLMVAVEGR1FRKWFAAAPDL1SYFIWD 1119

Db 755 HIVKNTRGPGLSRVITTKLDAPS-----QIEAK-----D 784

QY 1120 KTDVYNQKVFLSEAFVSVGYEYESCPDLILWEKRTTVLQGYE-----IDA 1165

Db 785 VTDT-----TALITWSKFLABIEGIELTYGPKDVPDRTTIDL 822

QY 1166 SKLGGWSLDKHHALNIQSGILHKNGENQFVSQQPPVIGSIMGNRRRS-----ISCPSC 1220

Db 823 SE-----DENQY-----SIGNLRPHTEYEYEVTLISR 849

QY 1221 NGLADGNKLLAPVALTCGSDGSLVVGDFNYIRRIIPSGNVN1LELRNKDFRHSHPAHK 1280

Db 850 RGMESDPA-----KEVFVTDLDAPRNLKRVSTQDNSITL---EWKFESHANIDN 895

QY 1281 XYLATDPMGSAVFLSDSNSRRVFKIKSTVVVKDLVKNSEVVAGTGDQCLPFDDTRCGDGG 1340

Db 896 YRIKFAPISGG-----DHTE1TVPKGNQA----- 919

QY 1341 KATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGI1ISTLLGSNDLTSARPLSCDSVMD 1400

Db 920 -TTRATLTGLRP-----GTEYGIGVTAVRQ-DRESAPATINAGTDLDPKOLEVSDPTE 971

QY 1401 ISQVLEWPTDLA-----INPMDSLYVLDNNVVLQISENHQVRIV 1441

Db 972 -TTLSLRWRRPVAKFDRYRLTYVSPSGKKNEMEI1PVDSTSFILRG---LDAGTEY1ISLV 1027

QY 1442 AGRPMHCQVP-----GIDHFLSKVAIHATLESATALAVSHNGV 1480

Db 1028 AEKGRHKSAPTITIKGSTEEEPBELNLSVSETGTGWDGFLTWTAAADGAYE-----NFV 1078

QY 1481 LYIAETDEKKINRIRQVTTSGEISLVAGAPS-----GDCCKNDANCDCFSGDDGYAKDAKL 1536

Db 1079 IQVQSDNPE-----ETWNITVPGGQHSVNVTLGLKANTPYNVTLYGVIRGYRTKPLY 1130

QY 1537 NTPSSLAVCADGELYVADL-----GNIRFIRKKNK---PF-----L 1570

Db 1131 VETTTGAHPEVGELTVSDITPESFNLSWTTTNGDFDAFIBIIDSNRLLPEMEFNISGNS 1190

QY 1571 NTQNMVELSSPIDQELYLFDITTKHLY-----TQSLPTGDYLYNFYTYTGDDITL 1620

Db 1191 RTAHSGLSPSTDFIVYLYGIS--HGFRTOAISAAATTEAPPEVDNLLVSDATPDGFRLT 1248

QY 1621 ITDNG-----NMVTVRRDSTGMP1LW1VP-----DGQVYVWTMGT 1656

Db 1249 WTADDGVFDSFVLKIRDTKRKSD--PLELIVPGHERTHDITGLKEGTEYEIE1LYGVSSGR 1306

QY 1657 NS-ALKSVTTQGHELAMMTYHGN-SGLLATKSNEN---GWTTFEYD1SFGRLTNVTFPT 1710

Db 1307 RSQPINSVAT-----TVVSGSPKGISFSDITENSARVSWTP-----PR 1343

QY 1711 GOVSSFRSDTDS1SVHVQVETSKDDVTITTNLSASGAFYTL1LQDQVRNSY1YGADGSLRL 1770

Db 1344 SRVDSYRVS-----YVPI1TGTEPNVTV1DGSKTR-----TKLVKLVP-----GVDYVNI 1388

QY 1771 LLANGMEVALQTEPHLLAGTVN1PTVGKRN--VTL1PD1NG1NL1V1EW1R-----QRK 1817

Db 1389 ISVKGE---ESEP--ISGILKTALDSPSGLVMMNITDSEALATWQPAIAAVDNIIVSYS 1443  
QY 1818 EQARGQVT--VFGRRLRVHNRNLLSLDFRVRTTEKIYDDHRKFTLRILYDQA--GRPSL 1873  
Db 1444 SEDEPEVTQMSGNIVE-YDLNGLRPATEYTLRVHAVKDAQKSETLSTQFTTGLDAPKDL 1502  
QY 1874 WSPSSRLNGVNVYSPGGYIAGIQGIMSERMEYDQAGRIITSRIFADGKTWSTYILEKSM 1933  
Db 1503 SATEVQSETAVITWRPP-----RAPVTDVLLTYESIDGRVKEVILDPETTSYTLTE--- 1553  
QY 1934 VLLHLSQRQYIFEPDKNDRLSSVTMPNVARQTLTIRSVGYRYNIYQPPEGNASVIQDFT 1993  
Db 1554 ---LSPSTQYTVKLQALSR----SMRSKMIQTVFTTTGLLY-----PYPKDCSQALLNGEV 1602  
QY 1994 EDGHLHLTFYLGTR 2008  
Db 1603 TSG--LYTIYLNQDR 1615

RESULT 12  
US-08-185-432-16  
; Sequence 16, Application US/08185432  
; Patent No. 5750652  
; GENERAL INFORMATION:  
; APPLICANT: Artavanis-Tsakonas, Spyridon  
; APPLICANT: Busseau, Isabelle  
; APPLICANT: Diederich, Robert J.  
; APPLICANT: Xu, Tian  
; APPLICANT: Matsuno, Kenji  
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND  
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/185,432  
; FILING DATE: 21-JAN-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2471 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-185-432-16

Query Match 3.0%; Score 440; DB 1; Length 2471;  
Best Local Similarity 30.1%; Pred. No. 3.8e-22;  
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;  
QY 562 SVDNCPN-CYNGDC---ISG-TCHCFLGLFLGPDGCRASCP-----VLCSGNGQY 607  
Db 795 NIDECASNPCLNQGTCTFDDISGYTCHCVLPYTGKNCQTVLAPCSPNPCENAAVCKESPNF 854  
QY 608 MKGRCLCHSGWKAECVDPTNQCIDVACSNHGTC--ITGT--CICNPGYKGSCEE--VD 661

Db 855 ESYTCLCAPGWQQRCTIDIDEICISKPCWNHGLCHNTQGSYMCBPPGFGSGMDCCEEDIDD 914  
QY 662 CMDPTCSGRGVCVRG----ECHCFVWGVTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714  
Db 915 CLANPCQNGSGMDGVNTFSCLCLPLGFTGDKCQDTMNECLSEPKNGGTCSDYVNSYTK 974  
QY 715 CDPSTWTHDC--SIEICA-ADCGHGVGVGG----TCRCEDGWMGAAC----DQACHPR 763  
Db 975 CQAGFDGVHCENNINECTESSCFNGGTCVDGINSFSCLPVGTGTSFCLHEINECSSHP- 1033  
QY 764 CAEHGTCRDG---KCECSPGWNGEHC-TIAHYLDRVVVKEGCPGLCNGNGRCTLDLNGWH 818  
Db 1034 CLNEGTCVDGLGTYRCSPLGYTGKNCQTVLNLCSR-----SPCKNKGTCVQKKAESQ 1086  
QY 819 CVCQLGWRGAGDTSMETACGDSKDNDDGLVDCMDPDCCLQPLCHINPLCLGSPN 874  
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLCHQHSVCINAGN 1131

RESULT 13  
US-08-083-590A-19  
; Sequence 19, Application US/08083590A  
; Patent No. 5786158  
; GENERAL INFORMATION:  
; APPLICANT: Artavanis-Tsakonas, S. et al.  
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods  
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And  
; TITLE OF INVENTION: Nucleic Acids  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/083,590A  
; FILING DATE: 25-JUN-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-015  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 8698864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2471 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-083-590A-19

Query Match 3.0%; Score 440; DB 1; Length 2471;  
Best Local Similarity 30.1%; Pred. No. 3.8e-22;  
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;  
QY 562 SVDNCPN-CYNGDC---ISG-TCHCFLGLFLGPDGCRASCP-----VLCSGNGQY 607  
Db 795 NIDECASNPCLNQGTCTFDDISGYTCHCVLPYTGKNCQTVLAPCSPNPCENAAVCKESPNF 854  
QY 608 MKGRCLCHSGWKAECVDPTNQCIDVACSNHGTC--ITGT--CICNPGYKGSCEE--VD 661

Db 855 ESYTCLCAPGWQORCTIDIDECISKPCNMHGLCHNTQGSYMCECPFGSGMDCEEDIDD 914  
QY 662 CMDPTCSGRGVCVRG----ECHCFVWGWTNCTPRATCLDQ-CSGHGTF--LPDTGLCS 714  
Db 915 CLANPCQGGSCMDGVNTFSCCLCPGFTGDKQOTDMNECLSEPCXNGGTCSDYVNSYTK 974  
QY 715 CDPSWTGHDC--SIEICA-ADCGGHGVCVGG---TCRCEDGWMGAAC----DQACHPR 763  
Db 975 CQAGFDGVHCENNINECTESSCFNGGTCVDGINSFSCCLCPVGFTGSFCLHEINECSSHP- 1033  
QY 764 CAEHGTCRDG---KCECSPGWNGEHC-TIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWH 818  
Db 1034 CLNEGTCVDGLGTYRCSCLPGYTGKNCQTLVNLCSR-----SPCKNKGTCVQKKAESQ 1086  
QY 819 CVCQLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPDCCCLQPLCHINPLCLGSPN 874  
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLCQHSVGCINAGN 1131

RESULT 14  
US-08-532-384-19  
; Sequence 19, Application US/08532384  
; Patent No. 6083904  
; GENERAL INFORMATION:  
; APPLICANT: Artavanis-Tsakonas, S. et al.  
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods  
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And  
; TITLE OF INVENTION: Nucleic Acids  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/532,384  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/083,590  
; FILING DATE: 25-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-015  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 8698864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2471 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-532-384-19

Query Match 3.0%; Score 440; DB 3; Length 2471;  
Best Local Similarity 30.1%; Pred. No. 3.8e-22;  
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;  
QY 562 SVDNCPN-CYNGDC---ISG-TCHCFLGFLGPDCCGRASCP-----VLCGNGQY 607  
Db 795 NIDECASNPCLNQGTCTFDDISGYTCHCVLPYTGKNCQTLVAPCSPNPCENAAVCKESPNF 854

QY 608 MKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTG--ITGT--CICNPGYKESCEE--VD 661  
Db 855 ESYTCLCAPGWQORCTIDIDECISKPCNMHGLCHNTQGSYMCECPFGSGMDCEEDIDD 914  
QY 662 CMDPTCSGRGVCVRG----ECHCFVWGWTNCTPRATCLDQ-CSGHGTF--LPDTGLCS 714  
Db 915 CLANPCQGGSCMDGVNTFSCCLCPGFTGDKQOTDMNECLSEPCXNGGTCSDYVNSYTK 974  
QY 715 CDPSWTGHDC--SIEICA-ADCGGHGVCVGG---TCRCEDGWMGAAC----DQACHPR 763  
Db 975 CQAGFDGVHCENNINECTESSCFNGGTCVDGINSFSCCLCPVGFTGSFCLHEINECSSHP- 1033  
QY 764 CAEHGTCRDG---KCECSPGWNGEHC-TIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWH 818  
Db 1034 CLNEGTCVDGLGTYRCSCLPGYTGKNCQTLVNLCSR-----SPCKNKGTCVQKKAESQ 1086  
QY 819 CVCQLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPDCCCLQPLCHINPLCLGSPN 874  
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLCQHSVGCINAGN 1131

RESULT 15  
US-08-899-232-1  
; Sequence 1, Application US/08899232  
; Patent No. 6436650  
; GENERAL INFORMATION:  
; APPLICANT: Artavanis-Tsakonas, Spyridon  
; APPLICANT: Qi, Huilin  
; TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON  
; FILE REFERENCE: 7326-046  
; CURRENT APPLICATION NUMBER: US/08/899,232  
; CURRENT FILING DATE: 1997-07-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2471  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-899-232-1

Query Match 3.0%; Score 440; DB 4; Length 2471;  
Best Local Similarity 30.1%; Pred. No. 3.8e-22;  
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;  
QY 562 SVDNCPN-CYNGDC---ISG-TCHCFLGFLGPDCCGRASCP-----VLCGNGQY 607  
Db 795 NIDECASNPCLNQGTCTFDDISGYTCHCVLPYTGKNCQTLVAPCSPNPCENAAVCKESPNF 854  
QY 608 MKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTG--ITGT--CICNPGYKESCEE--VD 661  
Db 855 ESYTCLCAPGWQORCTIDIDECISKPCNMHGLCHNTQGSYMCECPFGSGMDCEEDIDD 914  
QY 662 CMDPTCSGRGVCVRG----ECHCFVWGWTNCTPRATCLDQ-CSGHGTF--LPDTGLCS 714  
Db 915 CLANPCQGGSCMDGVNTFSCCLCPGFTGDKQOTDMNECLSEPCXNGGTCSDYVNSYTK 974  
QY 715 CDPSWTGHDC--SIEICA-ADCGGHGVCVGG---TCRCEDGWMGAAC----DQACHPR 763  
Db 975 CQAGFDGVHCENNINECTESSCFNGGTCVDGINSFSCCLCPVGFTGSFCLHEINECSSHP- 1033  
QY 764 CAEHGTCRDG---KCECSPGWNGEHC-TIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWH 818  
Db 1034 CLNEGTCVDGLGTYRCSCLPGYTGKNCQTLVNLCSR-----SPCKNKGTCVQKKAESQ 1086  
QY 819 CVCQLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPDCCCLQPLCHINPLCLGSPN 874  
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLCQHSVGCINAGN 1131

Search completed: June 24, 2004, 16:22:59  
Job time : 41 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:14:18 ; Search time 24 Seconds  
(without alignments)  
6007.594 Million cell updates/sec

Title: US-10-029-020-14  
Perfect score: 14887  
Sequence: 1 MDVKKPKPYRSLTRRDAER.....ELSDSANNIHFMRQSEMGRR 2769

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	652	4.4	1746	1	TENA_PIG
2	649.5	4.4	2201	1	TENA_HUMAN
3	644	4.3	1808	1	TENA_CHICK
4	597.5	4.0	4289	1	TENX_HUMAN
5	440	3.0	2471	1	NTC2_HUMAN
6	438.5	2.9	2703	1	NOTC_DROME
7	434	2.9	1213	1	JAG3_BRARE
8	429	2.9	2524	1	NOTC_XENLA
9	428	2.9	2556	1	NTC1_HUMAN
10	427.5	2.9	2531	1	NTC1_MOUSE
11	424	2.8	2470	1	NTC2_MOUSE
12	423.5	2.8	1064	1	FBP1_STRPU
13	422.5	2.8	2531	1	NTC1_RAT
14	420	2.8	2471	1	NTC2_RAT
15	419.5	2.8	1219	1	JAG1_RAT
16	413	2.8	2437	1	NTC1_BRARE
17	411.5	2.8	1218	1	JAG1_HUMAN
18	410.5	2.8	1242	1	JAG1_BRARE
19	405.5	2.7	1218	1	JAG1_MOUSE
20	404.5	2.7	1238	1	JAG2_HUMAN
21	400.5	2.7	1295	1	GLP1_CAEEL
22	397.5	2.7	2319	1	NTC3_RAT
23	395	2.7	1247	1	JAG2_MOUSE
24	392.5	2.6	2321	1	NTC3_HUMAN
25	391.5	2.6	1202	1	JAG2_RAT
26	387.5	2.6	1964	1	NTC4_MOUSE
27	386	2.6	2003	1	NTC4_HUMAN
28	384	2.6	2318	1	NTC3_MOUSE
29	383	2.6	1408	1	SERR_DROME
30	382.5	2.6	833	1	SRC2_MOUSE
31	382.5	2.6	2139	1	CRB_DROME
32	378.5	2.5	1376	1	CRBH_HUMAN
33	373.5	2.5	833	1	DL_DROME

34	369.5	2.5	830	1	SREC_HUMAN	Q14162	homo sapien
35	363	2.4	870	1	SRC2_HUMAN	Q96gp6	homo sapien
36	356.5	2.4	473	1	FP2_MVTGA	Q25464	mytilus gal
37	355	2.4	570	1	FBP3_STRPU	P49013	strongyloce
38	343.5	2.3	1429	1	LI12_CAEEL	P14585	caenorhabdi
39	342	2.3	723	1	DL11_HUMAN	Q00548	homo sapien
40	340.5	2.3	686	1	DL14_MOUSE	Q9ji71	mus musculu
41	340	2.3	714	1	DL11_RAT	P97677	rattus norv
42	337	2.3	1560	1	TENN_MOUSE	Q80z71	mus musculu
43	334.5	2.2	379	1	WIF1_HUMAN	Q9y5w5	homo sapien
44	331.5	2.2	379	1	WIF1_MOUSE	Q9wua1	mus musculu
45	327.5	2.2	378	1	WIF1_BRARE	Q9w6f9	brachydanio

ALIGNMENTS

RESULT 1  
TENA\_PIG  
ID TENA\_PIG STANDARD; PRT; 1746 AA.  
AC Q29116; P98142;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)  
DE (GMEM) (JI) (Mitochondious antigen) (Glioma-associated-extracellular  
DE matrix antigen) (GP 150-225) (Tenascin-C) (TN-C) (P230).  
GN TNC OR HXB.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS MAJOR; MINOR-1 AND MINOR-2).  
RC TISSUE=Submaxillary gland;  
RX MEDLINE=92104189; PubMed=1722152;  
RA Nishi T., Weinstein J., Gillespie W.M., Paulson J.C.;  
RT "Complete primary structure of porcine tenascin: detection of  
RT tenascin transcript in adult submaxillary glands.";  
RL Eur. J. Biochem. 202:643-648(1991).  
RN [2]  
RP SEQUENCE OF 813-825; 887-917; 998-1011; 1597-1608 AND 1719-1730.  
RC TISSUE=Fetal brain;  
RX MEDLINE=98158323; PubMed=9498558;  
RA Wakatsuki S., Ho S.H., Arioka M., Yamasaki M., Kitamoto K.;  
RT "Isolation and characterization of a 230 kDa protein (p230)  
RT specifically expressed in fetal brains: its involvement in neurite  
RT outgrowth from rat cerebral cortex neurons grown on monolayer of  
RT astrocytes.";  
RL J. Biochem. 122:1146-1152(1997).  
CC -!- FUNCTION: SAM (substrate-adhesion molecule) that appears to  
CC inhibit cell migration. May play a role in supporting the growth  
CC of epithelial tumors. Is a ligand for integrins alpha-8/beta-1,  
CC alpha-9/beta-1, alpha-v/beta-3 and alpha-v/beta-6.  
CC -!- FUNCTION: Plays a role during early brain development particularly  
CC in growth cone guidance. Involved in neurite outgrowth from  
CC cortical neurons grown on the monolayer of astrocytes.  
CC -!- SUBUNIT: Hexameric. A homotrimer may be formed in the triple  
CC coiled-coil region and may be stabilized by disulfide rings at  
CC both ends. Two of such half-hexabrachions may be disulfide linked  
CC within the central globule.  
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Comment=Isoforms are produced in a tissue-and time-specific  
CC manner during development;  
CC Name=Minor-2;  
CC IsoId=Q29116-1; Sequence=Displayed;  
CC Name=Major;  
CC IsoId=Q29116-2; Sequence=VSP\_001416;  
CC Name=Minor-1;  
CC IsoId=Q29116-3; Sequence=VSP\_001417;  
CC -!- TISSUE SPECIFICITY: Submaxillary glands and brain.



Db 535 DCHGRGRC---VNG-QCVCHGFTGKDCGQRRCPGDCHGQRCVDGQCVCHEGFTGLDCG 590  
QY 854 DPDC-----CLQPLCHINPLCLG-----SPNPLDIIQETQVPVSQQNLHSFYDR- 897  
Db 591 QRSQPNDCSNWQCVCVSGRCICNEGYSGEDCSQVSP-PKDLI---VTEVTEETVNLAWDNE 646  
QY 898 ---IKFLVGRDSTH-----IIPGENPFDGGHACVIRGOVMTSDGTPLGVN-----I 941  
Db 647 MRVTEYLIVYTPTHEDEGLEMQFRVPGD-----QSTTIRELEPGVEYFIRVF 693  
QY 942 SFVNNPLFGYTTISRQDSFDLVNTGGISIIILRFRAPFITQEHTLWLWDPRFFVMETIIM 1001  
Db 694 AILENKK-SIPVSARVATY-LPTPEG---LKFKSIK-ETSEVEVWDPLDIAFETWEIIF 746  
QY 1002 RHEENEIPSCDLSNFARENPNVPSPLTSFASSCAEKGPVPEIQALQEBEISISGCKMRL 1061  
Db 747 RNMNKEDEGEITKSLRPE-----TTYRQTGLAPG-----QBEYELS---LHI 785  
QY 1062 SYLSSRTPGYKSVLRISLTHPTTIPFNLKMKVHLMVAVEGRLEFRKWFAAAPDLSYFIWDKT 1121  
Db 786 VKNTRGPGGLKRVTTTTRLDAPS-----QIEAKDVTDTTALITWFKPLAID----- 831  
QY 1122 DVYNQKVFLSEAFVSVGYEYESCPDLILWEKRTVLQGYEIDASKLGWSLDKHHALNI 1181  
Db 832 -----GIELTYGIKDVPG-----DRATIDLTHEENQYSIGNLKPDEYEVSL 873  
QY 1182 QSGILHKGNGENQFVSQOPPVIGSIMGNRRRSISPCSCNGLADGNKLLAPVALTCGSDG 1241  
Db 874 IS-----RR-----ADMSSNPAKETFTTGLDA 895  
QY 1242 SLVVGDFNYIRIFPSPGNVTNILELRNKDFRHSHPAKHYLATDPMGAVFLSDSNSRR 1301  
Db 896 P-----RNLRRISQTDNSITL-----EWRNGKAAADTYRIKYAPISGG----- 933  
QY 1302 VFKIKSTVVVKDLVKNSEVVAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLI 1361  
Db 934 -----DHAEVEVPRSPQ-----TTTKATLTGLRPGT--EYGI- 963  
QY 1362 YFVDGTMIRRIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLA-----IN 1415  
Db 964 ----GVS AVKGDKESDPATINAATDLDPKDFRVSELKE-SSLTLLWRTPLAKFDYRLN 1018  
QY 1416 -----PMDNSLYVLDDNNVVLQISENHQVRIVAGRPMHCQVPG----- 1452  
Db 1019 YGLPSGQPVVOLPRNATSYILRG--LEPGQEYITILLTAEKGRHKSKPARVKASTAGEP 1075  
QY 1453 -----IDHFLSKVAIHATLESATALAVSHNGVLVYIAETDEKKNRI-RQVTT 1500  
Db 1076 EIGNLSVSDITPESFSLSWTATEGAFETFTTIEIDSNRFL---ETMEYNISGAERTAHIS 1132  
QY 1501 G-----EISLVAGAPSGCDCKNDANCDCFGSDGDAYAKDAKLNTPSSLAVCADGE----- 1549  
Db 1133 GLRPGNDFIVYLSGLAPGIQTK-----PISATATTEAEPEVDN 1170  
QY 1550 LYVADLGNIRIRFIRKNKPPFLNTQNMVELSSPIDQELY-----LFDTTGKHLYTQS LPT 1603  
Db 1171 LLVSDA-----TPDGFRLSWTADBGVFDSPVLKIRDTK-----KQSEPL 1209  
QY 1604 GDLYNFTYTGDGDTITLTDNNGMVMNVRRDSTGMPLWLVPDGVVYVWMTGNTSALKSV 1663  
Db 1210 -----EITLLASERTDITGLREATEYEI-----ELYGISGKRS----- 1244  
QY 1664 TQGHELAMTYHGNSGLLATKSNENGWTTTFYEYDSFGRLTNVTFTPTGOVSSFRSDTSS 1723  
Db 1245 -----QPVSAIATTAMGSPKEITFSDITENSATVSMVPPTAQVESFR-----I 1287  
QY 1724 VHVQVETSSKDDVTITTNLSASGAFYTLTLLQDQVNSYYIGADGSLRLILLANGMEVALQTE 1783  
Db 1288 TYVPITGGAPSVVTVDGTKTQTRLRLLL-----PGVEYLVSVIAVKGFE---ESE 1334  
QY 1784 PHLLAGTVNPTVGKRVNPLPIDNGLNLVEWRQKQEQARGQVTVFGRRRLRVHNRNLLSLDF 1843

Db 1335 P--VSGT-----LTTALDGPSPGLVT----- 1352  
QY 1844 DRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPS-SRLNGVNVVTPSPGGYIAGIQGIMS 1902  
Db 1353 ANITDSEAL-----AMWQPAIAPVDHYVISYT-GDRVPEITRIVSG 1392  
QY 1903 ERMEY-----DQAGRITSRIFAD-GKTWSYTYLEKSMVLLHLSQRQYIFEFDKNDRLSSV 1956  
Db 1393 NTVYALTNLPEATEYTLRIFAEGPKQKSSITTK-----FTTDL-DSPRDL 1438  
QY 1957 TMPNVARQT-LETIR-----SVGYRNIIYQPPEGNASVI-----QDFTEDGHLHLHTFYLG 2005  
Db 1439 TATEVQSETALLTWRPPRASVTGYLLVYESVDGTLKEVVVGPETTTSYLSGLSPSTHYTA 1498  
QY 2006 TGRVVIYKYGKL-SKLAETLYDITTKVSFTY-----DETAGMLKTINLQNE 2049  
Db 1499 ---RIQALNGPLRSKMSQVFTTIGLLYPFRDCSQAMLNGDTSGLYTIYVNNND 1550

RESULT 2

TENA\_HUMAN  
ID TENA\_HUMAN STANDARD; PRT; 2201 AA.  
AC P24821; Q14583; Q15567;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)  
DE (GMEM) (JI) (Mitochondinous antigen) (Glioma-associated-extracellular  
DE matrix antigen) (GP 150-225) (Tenascin-C) (TN-C).  
GN TNC OR HXB.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-32.  
RC TISSUE=Fetal brain, and Melanoma;  
RX MEDLINE=91187670; PubMed=1707164;  
RA Siri A., Carnemolla B., Saginati M., Leprini A., Casari G.,  
RA Baralle F., Zardi L.;  
RT "Human tenascin: primary structure, pre-mRNA splicing patterns and  
RT localization of the epitopes recognized by two monoclonal  
RT antibodies.";  
RL Nucleic Acids Res. 19:525-531(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95155442; PubMed=7531707;  
RA Gherzi R., Carnemolla B., Siri A., Ponassi M., Balza E., Zardi L.;  
RT "Human tenascin gene. Structure of the 5'-region, identification, and  
RT characterization of the transcription regulatory sequences.";  
RL J. Biol. Chem. 270:3429-3434(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91131572; PubMed=1704365;  
RA Nies D.E., Hemesath T.J., Kim J.H., Gulcher J.R., Stefansson K.;  
RT "The complete cDNA sequence of human hexabrachion (Tenascin). A  
RT multidomain protein containing unique epidermal growth factor  
RT repeats.";  
RL J. Biol. Chem. 266:2818-2823(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92052108; PubMed=1719530;  
RA Gulcher J.R., Nies D.E., Alexakos M.J., Ravikant N.A., Sturgill M.E.,  
RA Marton L.S., Stefansson K.;  
RT "Structure of the human hexabrachion (tenascin) gene.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9438-9442(1991).  
RN [5]  
RP SEQUENCE OF 431-2055 FROM N.A. (ISOFORM 6).  
RC TISSUE=Glioblastoma;  
RX MEDLINE=89160821; PubMed=2466295;  
RA Gulcher J.R., Nies D.E., Marton L.S., Stefansson K.;  
RT "An alternatively spliced region of the human hexabrachion contains a  
RT repeat of potential N-glycosylation sites.";







FT	CONFLICT	182	182	FTid=VSP_001411.
FT	CONFLICT	563	571	W -> R (IN REF. 2).
FT	CONFLICT	598	598	SCPNDNNV -> PAPMTATTW (IN REF. 3).
FT	CONFLICT	838	838	E -> G (IN REF. 3).
FT	CONFLICT	886	886	T -> TEY (IN REF. 3).
FT	CONFLICT	886	886	N -> F (IN REF. 3).
SQ	SEQUENCE	1808	AA; 198858	MW; B924A06CF9EFD6DE CRC64;
Query Match				
Best Local Similarity 4.3%; Score 644; DB 1; Length 1808;				
Matches 334; Conservative 219; Mismatches 566; Indels 510; Gaps 71;				
Qy	543	FYNDGKSEVVSFLTTAIESVDNCPNSCYNGGDCISGTHCHFLGFLGPDCCGRASCPVLCS	602	
Db	332	FCEEGYTGDCGELT-----CPNNCNGNGRCENGCLVCVCHGEFVGDDCSKRCPKDCN	383	
Qy	603	GNQYMKGRCLCHSGWKGAECD--VPTNQCIDVACSNHGTCTITGTICIONPGYKGESCEE	659	
Db	384	NRGHCVDGRVCVCHGYLGEDCGELRCPND-----CHNRGRACINGQCVDEGFIGEDCGE	437	
Qy	660	VDCMDPTCSGRGVCVRGECHCFVGGTNGTCTPRATCLDQCSCGHGTFPLDTGLCSCDPSW	719	
Db	438	LRCPN-DCHNRGRVCNGQCECHGEFVIGEDCGELR--CPNDCNSHGRCV--NGQCVCDEGY	492	
Qy	720	TGHDCSIEICAADCGHGVCGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSP	779	
Db	493	TGEDCGELRCPNDCHNRGRVCVEGRVCVCDNGFMGEDCGELSCPNDCHQHGRCVDGRCVCHE	552	
Qy	780	GWNGEHCCTIAHYLDRVVKEGCPGLCNGNGRCRTLDLNGWHVCVQLGWRGAGCDTSMETACG	839	
Db	553	GFTGEDCR-----ERSCPNDCNNVGR-----VEG-RCVCEEGYMG-----	587	
Qy	840	DSKNDGDGLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSOQNLHSFYDRI-	898	
Db	588	-----IDCSD-----VSP-----PTELTVTNVTDKTV---NLEWKHENLV	619	
Qy	899	KFLV-----GRDSTHIIIGE-----NPFDDGGHACVIRGOVMTSDGTPLVGVNISF	943	
Db	620	NEYLVTYVPTSSGGLDLQFTVPGNQTSATHELPEGVVEYFIRVFAILKNKXSI-----	672	
Qy	944	VNNPLFGYTIHQDGSFDLVTNGGSIILRFRERAPFITQHTLWLPWDRF---FVMTIIL	1000	
Db	673	-----PVSARVATYLPAPEG-----LKFKS-----VRETSVQVVEWDPLSISFDGWELV	715	
Qy	1001	MRHEENEIPSCDL-SNFARPNNPVSPSPLTSFASCAEKGPVPEIQALQEEISISGCKM	1059	
Db	716	FRNMQKKODNGDITSSLKRPE-----TSYMQPGLAPG-----QQYNVS---L	754	
Qy	1060	RLSYLSSRTPGYKSVLRISLTHPTIPFNLMKVHLMVAVEGRLFRKWFAAAPDLSSYFIWD	1119	
Db	755	HIVKNTRGPGLSRVITTKLDAPS-----QIEAK-----D	784	
Qy	1120	KTDVYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTVLQGYE-----IDA	1165	
Db	785	VTDT-----TALITWSKPLABIEGIELTYGPKDVPGRDRTIDL	822	
Qy	1166	SKLGWSLDKHHALNIQSGILHKNGENQFVSQQPPVIGSIMGNRRRS---ISCPSCNG	1222	
Db	823	SE-----DENQY-----SIGNLRPHTEYEVTLLISRRG	849	
Qy	1223	LADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFPSGNVTNILELRNKDFRHSHPAHKYV	1282	
Db	850	DMESDPA-----KEVFVTDLDAPRLKRVSTQDNSITL---EWKNSHANIDNYR	895	
Qy	1283	LATDPMGAVFLSDSNSRRVFKIKSTVVVKDLVKNSEVVAGTGDCQLPLFDDTRCGDGKA	1342	
Db	896	IKFAPISGG-----DHTELTVPKGNQA-----T	918	
Qy	1343	TEATLTNPRGITYDKFGLIYFVDGTMIRRIDQNGIISILLGSNDLTARSPLSCDSVMDIS	1402	
Db	919	TRATLTGLRP-----GTEYICIGVTAVRQ-DRESAPATINAGTDLDPKDLVSDPTE-T	970	
Qy	1403	QVRLEWPTDLA-----INPMDNSLVLDNNNVVLQISENHQVRIVAG	1443	

RP SEQUENCE OF 1-747 AND 1687-1944 FROM N.A.  
RC TISSUE=Leukocyte;  
RX MEDLINE=93300909; PubMed=7686164;  
RA Bristow J., Tee M.K., Gitelman S.E., Mellon S.H., Miller W.L.;  
RT "Tenascin-X: a novel extracellular matrix protein encoded by the human  
RL XB gene overlapping P450c21B.";  
RN J. Cell Biol. 122:265-278(1993).  
[3]  
RP SEQUENCE FROM N.A. (ISOFORM XB-SHORT).  
RC TISSUE=Adrenal gland;  
RX MEDLINE=96015044; PubMed=8530023;  
RA Tee M.K., Thomson A.A., Bristow J., Miller W.L.;  
RT "Sequences promoting the transcription of the human XA gene  
RT overlapping P450c21A correctly predict the presence of a novel,  
RT adrenal-specific, truncated form of tenascin-X.";  
RL Genomics 28:171-178(1995).  
[4]  
RP SEQUENCE OF 1-23 FROM N.A.  
RC TISSUE=Fetal adrenal gland;  
RX MEDLINE=97081760; PubMed=8923003;  
RA Speek M., Barry F., Miller W.L.;  
RT "Alternate promoters and alternate splicing of human tenascin-X, a  
RT gene with 5' and 3' ends buried in other genes.";  
RL Hum. Mol. Genet. 5:1749-1758(1996).  
[5]  
RP SEQUENCE OF 3470-4289 FROM N.A.  
RX MEDLINE=89367293; PubMed=2475872;  
RA Morel Y., Bristow J., Gitelman S.E., Miller W.L.;  
RT "Transcript encoded on the opposite strand of the human steroid 21-  
RT hydroxylase/complement component C4 gene locus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:6582-6586(1989).  
[6]  
RP DISEASE.  
RX MEDLINE=21468843; PubMed=11642233;  
RA Schalkwijk J., Zweers M.C., Steijlen P.M., Dean W.B., Taylor G.,  
RA van Vlijmen I.M., van Haren B., Miller W.L., Bristow J.;  
RT "A recessive form of the Ehlers-Danlos syndrome caused by tenascin-X  
RT deficiency.";  
RL New Engl. J. Med. 345:1167-1175(2001).  
CC -!- FUNCTION: Appears to mediate interactions between cells and the  
CC extracellular matrix. Substrate-adhesion molecule that appears to  
CC inhibit cell migration. May play a role in supporting the growth  
CC of epithelial tumors.  
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Comment=Additional isoforms seem to exist;  
CC Name=XB;  
CC IsoId=P22105-1; Sequence=Displayed;  
CC Name=XB-short;  
CC IsoId=P22105-2; Sequence=VSP\_001418;  
CC -!- TISSUE SPECIFICITY: Highly expressed in fetal adrenal, in fetal  
CC testis, fetal smooth, striated and cardiac muscle. Isoform XB-  
CC short is only expressed in the adrenal gland.  
CC -!- DISEASE: Association with congenital adrenal hyperplasia.  
CC -!- DISEASE: Defects in TNXB are the cause of Ehlers-Danlos-like  
CC syndrome [MIM:606408]. This clinically distinct form of Ehlers-  
CC Danlos syndrome is characterized by hyperextensible skin,  
CC hypermobile joints, and tissue fragility, but it lacks atrophic  
CC scars and delayed wound healing. Inheritance is autosomal  
CC recessive.  
CC -!- SIMILARITY: Contains 19 EGF-like domains.  
CC -!- SIMILARITY: Contains 32 fibronectin type III domains.  
CC -!- SIMILARITY: Contains 1 fibronectin C-terminal domain.  
CC -!- CAUTION: There are two genes for TN-X: TNXA and TNXB. TNXA is a  
CC partial gene which can sometimes recombine with TNXB.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC -----  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U89337; AAB47488.1; -.  
DR EMBL; AF019413; AAB67981.1; -.  
DR EMBL; X71923; CAA50739.1; -.  
DR EMBL; Y13782; CAA74109.1; -.  
DR EMBL; Y13783; CAA74110.1; -.  
DR EMBL; U24488; AAB41287.1; -.  
DR EMBL; U52696; AAC50889.1; -.  
DR EMBL; M25813; AAA35884.1; -.  
DR PIR; A40701; A40701.  
DR HSSP; P02671; 1FZD.  
DR Genew; HGNC:11976; TNXB.  
DR MIM; 600985; -.  
DR MIM; 606408; -.  
DR GO; GO:0005578; C:extracellular matrix; NAS.  
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR002181; Fibrinogen C.  
DR InterPro; IPR008957; FN\_III-like.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR Pfam; PF00008; EGF; 8.  
DR Pfam; PF00147; fibrinogen\_C; 1.  
DR Pfam; PF00041; fn3; 33.  
DR PRINTS; PR00011; EGFLAMININ.  
DR SMART; SM00181; EGF; 7.  
DR SMART; SM00186; FBG; 1.  
DR SMART; SM00060; FN3; 30.  
DR PROSITE; PS00022; EGF\_1; 18.  
DR PROSITE; PS01186; EGF\_2; 19.  
DR PROSITE; PS50026; EGF\_3; 8.  
DR PROSITE; PS00514; FIBRIN AG C DOMAIN; 1.  
KW Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;  
KW Extracellular matrix; Alternative splicing; Signal;  
KW Ehlers-Danlos syndrome.  
FT SIGNAL 1 23  
FT CHAIN 24 4289  
FT DOMAIN 156 168  
FT DOMAIN 183 213  
FT DOMAIN 214 244  
FT DOMAIN 245 275  
FT DOMAIN 276 306  
FT DOMAIN 307 337  
FT DOMAIN 338 368  
FT DOMAIN 369 399  
FT DOMAIN 400 430  
FT DOMAIN 431 461  
FT DOMAIN 462 492  
FT DOMAIN 493 523  
FT DOMAIN 524 554  
FT DOMAIN 555 585  
FT DOMAIN 586 616  
FT DOMAIN 617 647  
FT DOMAIN 648 679  
FT DOMAIN 684 714  
FT DOMAIN 715 746  
FT DOMAIN 792 872  
FT DOMAIN 901 922  
FT DOMAIN 941 1021  
FT DOMAIN 1047 1127  
FT DOMAIN 1149 1226  
FT DOMAIN 1246 1327  
FT DOMAIN 1348 1429  
FT DOMAIN 1459 1540  
FT DOMAIN 1561 1642  
FT DOMAIN 1659 1736  
FT DOMAIN 1756 1836  
FT DOMAIN 1856 1939  
FT DOMAIN 1962 2039  
FT DOMAIN 2069 2150  
FT DOMAIN 2167 2248  
FT TENASCIN X.  
FT EGF-LIKE 1 (INCOMPLETE).  
FT EGF-LIKE 2.  
FT EGF-LIKE 3.  
FT EGF-LIKE 4.  
FT EGF-LIKE 5.  
FT EGF-LIKE 6.  
FT EGF-LIKE 7.  
FT EGF-LIKE 8.  
FT EGF-LIKE 9.  
FT EGF-LIKE 10.  
FT EGF-LIKE 11.  
FT EGF-LIKE 12.  
FT EGF-LIKE 13.  
FT EGF-LIKE 14.  
FT EGF-LIKE 15.  
FT EGF-LIKE 16.  
FT EGF-LIKE 17.  
FT EGF-LIKE 18.  
FT EGF-LIKE 19.  
FT FIBRONECTIN TYPE-III 1.  
FT COILED COIL (POTENTIAL).  
FT FIBRONECTIN TYPE-III 2.  
FT FIBRONECTIN TYPE-III 3.  
FT FIBRONECTIN TYPE-III 4.  
FT FIBRONECTIN TYPE-III 5.  
FT FIBRONECTIN TYPE-III 6.  
FT FIBRONECTIN TYPE-III 7.  
FT FIBRONECTIN TYPE-III 8.  
FT FIBRONECTIN TYPE-III 9.  
FT FIBRONECTIN TYPE-III 10.  
FT FIBRONECTIN TYPE-III 11.  
FT FIBRONECTIN TYPE-III 12.  
FT FIBRONECTIN TYPE-III 13.  
FT FIBRONECTIN TYPE-III 14.

FT	DOMAIN	2266	2347	FIBRONECTIN TYPE-III 15.
FT	DOMAIN	2365	2446	FIBRONECTIN TYPE-III 16.
FT	DOMAIN	2471	2552	FIBRONECTIN TYPE-III 17.
FT	DOMAIN	2582	2663	FIBRONECTIN TYPE-III 18.
FT	DOMAIN	2688	2769	FIBRONECTIN TYPE-III 19.
FT	DOMAIN	2794	2875	FIBRONECTIN TYPE-III 20.
FT	DOMAIN	2889	2972	FIBRONECTIN TYPE-III 21.
FT	DOMAIN	2997	3078	FIBRONECTIN TYPE-III 22.
FT	DOMAIN	3105	3186	FIBRONECTIN TYPE-III 23.
FT	DOMAIN	3211	3292	FIBRONECTIN TYPE-III 24.
FT	DOMAIN	3307	3384	FIBRONECTIN TYPE-III 25.
FT	DOMAIN	3399	3481	FIBRONECTIN TYPE-III 26.
FT	DOMAIN	3494	3575	FIBRONECTIN TYPE-III 27.
FT	DOMAIN	3601	3682	FIBRONECTIN TYPE-III 28.
FT	DOMAIN	3699	3787	FIBRONECTIN TYPE-III 29.
FT	DOMAIN	3801	3879	FIBRONECTIN TYPE-III 30.
FT	DOMAIN	3890	3971	FIBRONECTIN TYPE-III 31.
FT	DOMAIN	3978	4059	FIBRONECTIN TYPE-III 32.
FT	DOMAIN	4071	4289	FIBRINOGEN C-TERMINAL.
FT	SITE	1748	1750	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	187	197	BY SIMILARITY.
FT	DISULFID	191	202	BY SIMILARITY.
FT	DISULFID	204	213	BY SIMILARITY.
FT	DISULFID	218	228	BY SIMILARITY.
FT	DISULFID	222	233	BY SIMILARITY.
FT	DISULFID	235	244	BY SIMILARITY.
FT	DISULFID	249	259	BY SIMILARITY.
FT	DISULFID	253	264	BY SIMILARITY.
FT	DISULFID	266	275	BY SIMILARITY.
FT	DISULFID	280	290	BY SIMILARITY.
FT	DISULFID	284	295	BY SIMILARITY.
FT	DISULFID	297	306	BY SIMILARITY.
FT	DISULFID	311	321	BY SIMILARITY.
FT	DISULFID	315	326	BY SIMILARITY.
FT	DISULFID	328	337	BY SIMILARITY.
Query Match 4.0%; Score 597.5; DB 1; Length 4289;				
Best Local Similarity 20.7%; Pred. No. 1.2e-25;				
Matches 325; Conservative 158; Mismatches 515; Indels 571; Gaps 66;				
QY	562	SVDNCPSCYNGDCISGTCHCFLGFLGPDGGRASCPVLCSGNGQYMKGRCLCHSGWKA	621	
Db	369	STRTCPDRGRGRCEDEICIDTGYSGDDCGVSRCPGDCNQGRCEDCRCVCWPYGTGT	428	
QY	622	ECDVPTNQCIDVACSNGHTCITGTICINPGYKGESCEEVDMDPTCSGEGVCVRGECHCF	681	
Db	429	DCG---SRACPRDCRGRCEGVCVGNAGYSGEDCGVRSC-PGDCRGRGRCESGRCMCW	484	
QY	682	VGWGGTNCET-----PRATCLD-----	712	
Db	485	PYTGTRDCGTRACPGDCRGRGRVCVGNPGETGEDCGSRRRCPGDCRCHG--LCEDGV	542	
QY	713	CSCDPSWTGHDCSIIEICAADCGGHGVCVGGTCRCEDGWMGAACDQACHPRCAEHGTCRD	772	
Db	543	CVCDAGYSGEDCSTRCPGGCRGRGQCLDGRVCVCEDDGYSGEDCGVRQCPNDCSQHGVQCD	602	
QY	773	GKCECSPGNGEHCTTAHYLDRVVKEGCPGLCNGNGRCRTLNLNGWHCVCOLGWRGAGCDT	832	
Db	603	GVCICWEGYVSEDCSI-----RTCPSNCHGRGRC-----EEGRCLCDPGYTGPTCAT	649	
QY	833	SMETACGDSKNDGDGLVDCMDPDCCLOP--LCHIN-----	877	
Db	650	RMCPA-----DCRGRGRVCQGVCLCHVGYGGEDCGQEPPASACFGCGGPRE	696	
QY	878	IIQ-----	890	
Db	697	LCRAGQCVCEGFRGPDCAIQTCPGDCRGRGECHDGCVCVKDGYAGEDCGEARVPSSA--	754	
QY	891	LHSFYDRIKFLVGRDSTHIIPIGENPFDGGHACVIRGQWMTSDGTPLVGVNISFVNNPLFG	950	
Db	755	--SAYDQ-----RGLAPQE-----YQVTVRALRGTSWGLPASKTITTWIDGP---	795	
QY	951	YTIHQDGSFDLVTNGGSIILRPERAFITQEHTLWLPW-----DRFFV-----	995	

Db	796	---QDLRVAVT-----PITLEGLWRPQAEVDRFVVSVYSAGNQ	832	
QY	996	-----METII-----MRHEENEIIPSCDLSN	1015	
Db	833	RVRLEVPPEADGTLTDLMPGVEYVTVTAERGRAVSYPASVRANTEEREESP-----	886	
QY	1016	FARPNPVSPSPLTSFASSCAEKGPVPEIQALQOEIISISGCKMRLSYLSSRTPGYKSVL	1075	
Db	887	---PRPSLSQPPRRPWGNLTAELSRFRGTVDLERHLRAHYPLR-----ANQT--YTSVA	937	
QY	1076	RISLTHPTIPFNLKVH-----LMVAVEG-----RLFRKWFAAAP-----	1110	
Db	938	R-----HIHEYLQROVLGSSADGALLVSLDGLRGQFERVVLWRPQPPAEGP	984	
QY	1111	--DLSYFYFIWDKTDVYNQKVFGLSFAFVSUGVEYVESCPDDLILWEKRTTVLQGYEIDASKL	1168	
Db	985	GGELT-----VPGTTRTV-----SLPDL---RPGTT---YHVEVHGV	1015	
QY	1169	GGWSLDKHHALNIQSGILHKGNGENQFVSQQPPVIGSIMNGR-----	1211	
Db	1016	RAGQTSKSYAFITTTGPSTTGAQAPLLQORPQELGELRVLGRDETGRVVTWTAQPDTF	1075	
QY	1212	-----RRSISCP-----SCNGLADGNKLLAPVALTCG	1238	
Db	1076	AYFQLMRMRVPEGPGAHEEVLPGDVRQALVPPPPCTPYELSLHGVPFGKPSDPIIYQGI	1135	
QY	1239	SD-----GSLVY---GDFNYIRRIFFPSGNVTN-ILELRNKDFRHSHPA	1278	
Db	1136	MDKDEEKPGRSGPPRLGELTVTDRTSDSLLLRWTVPGEFDSFVIQYKDRDGOQVWPV	1195	
QY	1279	HKYVLATDPMGSAVFLS-DSNSRRRVFKIKSTVVVKDLVKNSEVWAGTGDCQLPFPDTRCG	1337	
Db	1196	E-----GPQSAVITSLDPGRKYKFVLYGFVGK---RHGPLVAEA--KILPOSDDSPG	1244	
QY	1338	DGGKATEATLTNPRGITVDKFLIYFV---DGTWIRRIDQNG-----II	1378	
Db	1245	TPHLGNLWVTDp---TPDSLHLSWTVPGEQFDTFMVQYRDRDGRPQVPVPEGERSFVV	1301	
QY	1379	S-----TLGSDNLTSAARPLSCDSV-----MDI	1401	
Db	1302	SSLDPDHKYRFTLFGIANKKRYGPLTADGTTAPERKEBPPEFLEQPLLGELTVGTVP	1361	
QY	1402	SQVRLEWPTDLAINPMDNSLYVLDDNNVLQISE---NHQVRIVAGRPMHCQVPGIDHFL	1458	
Db	1362	DSLRLSW--TVAQGPFD-----SFMVQYKDAQQOPQAVPVAGDENEVTVPGLDPDRK	1411	
QY	1459	SKVAIHATL-----ESATALAVSHNGV-----LYIAETDEKKINRIRQVTTSG	1501	
Db	1412	YKMDLYGLRGRQRVGPESVVAKTAPQEDVDETSPTELGTAEAPESPPEPL--LGELTVTG	1469	
QY	1502	---EISLVAGAPSGCDCKNDANCDGSGDDGYAKADAKLNTPSSLAVCADGELYVADLGN	1557	
Db	1470	SSPDSLFLWTVPQG-----SFDSFTVQYKDRDGRPRAVRVGKSESEVTVGG	1516	
QY	1558	IRIRFIRKNKPPFLNTQNMVE-----LSSPIDQ---ELYLFDT	1591	
Db	1517	LE---PGHKYKMHLYGLHEGQRVGPVSAVGVTAPQOEBETPPATESPLEPRLGELTVTDV	1572	
QY	1592	TGXHL-YTQSLPTGDYLYNFTYTGDDITLITDNNG--NMVNVRRDSTGMPLWLVP---	1645	
Db	1573	TPNSVGLSWTVPEGQF-----DSFIVQYKDKGQPOVVPVAADQREVTVYNLEPERK	1624	
QY	1646	-----DGQVYVVTMTGINSALKSVTTQGHELAMMTYHGNSGLLATKSNENGWTTFFE	1696	
Db	1625	YKMNMYGLHDGQ---RMGFLSVV-IVTAPATEASKPPELPRLGELT-----VTDIT	1671	
QY	1697	YDSFGRLTNVTFTPTGOVSSF---RSDTDSVHVQVETSSKDDVTITTNLSASGAFYTL-	1752	
Db	1672	PDSVG--LSWTVPGEFDSFVVQYKDRDGOQVVPVVAADQREVTVYKFKLLF	1728	
QY	1753	--QDQVRNS	1759	

Db 1729 GIQDKRRS 1737

RESULT 5

NTC2\_HUMAN

ID NTC2\_HUMAN STANDARD; PRT; 2471 AA.

AC Q04721; Q99734; Q9H240;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (hN2).

GN NOTCH2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Blaumueller C.M., Mann R.S.;

RT "Complete human notch 2 (hN2) cDNA sequence.";

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Breast tumor;

RA Correa R.G., Camargo A.A., Moreira E.S., Simpson A.J.G.;

RT "Human Notch2, a novel member of cell-fate determining NOTCH family.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 967-1229 FROM N.A.

RC TISSUE=T-cell;

RA Lemasson I., Devaux C., Mesnard J.M.;

RT "Partial sequence of EGF-like repeat domain of human Notch2 mRNA.";

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 1810-2447 FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=93265135; PubMed=1303260;

RA Stifani S., Blaumueller C.M., Redhead N.J., Hill R.E.,

RA Artavanis-Tsakonas S.;

RT "Human homologs of a Drosophila enhancer of split gene product define a novel family of nuclear proteins.";

RL Nat. Genet. 2:119-127(1992).

RN [5]

RP POST-TRANSLATIONAL PROCESSING.

RX MEDLINE=97386453; PubMed=9244302;

RA Blaumueller C.M., Qi H., Zagouras P., Artavanis-Tsakonas S.;

RT "Intracellular cleavage of Notch leads to a heterodimeric receptor on the plasma membrane.";

RL Cell 90:281-291(1997).

RN [6]

RP IDENTIFICATION OF LIGANDS.

RX MEDLINE=99180765; PubMed=10079256;

RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,

RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;

RT "Human ligands of the Notch receptor.";

RL Am. J. Pathol. 154:785-794(1999).

CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands Jagged1, Jagged2 and Delta1 to regulate cell-fate determination. Upon ligand activation through the released notch intracellular domain (NICD) it forms a transcriptional activator complex with RBP-J kappa and activates genes of the enhancer of split locus. Affects the implementation of differentiation, proliferation and apoptotic programs (By similarity).

CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-terminal fragment N(EC) which are probably linked by disulfide bonds (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following proteolytical processing NICD is translocated to the nucleus.

CC -!- TISSUE SPECIFICITY: Expressed in the brain, heart, kidney, lung, skeletal muscle and liver.

CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form

CC which is proteolytically cleaved by a furin-like convertase in the trans-Golgi network before it reaches the plasma membrane to yield an active, ligand-accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(EC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (TACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NICD) from the membrane (By similarity).

CC -!- PTM: Phosphorylated (By similarity).

CC -!- SIMILARITY: Belongs to the NOTCH family.

CC -!- SIMILARITY: Contains 35 EGF-like domains.

CC -!- SIMILARITY: Contains 2 Lin/Notch repeats.

CC -!- SIMILARITY: Contains 6 ANK repeats.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC -----

CC EMBL; AF308601; AAA36377.2; -

CC EMBL; AF315356; AAG37073.1; -

CC EMBL; U77493; AAB19224.1; -

CC HSSP; P00740; 1EDM.

CC Genew; HGNC:7882; NOTCH2.

CC MIM; 600275; -

CC InterPro; IPR002110; ANK.

CC InterPro; IPR000152; Asx\_hydroxyl\_S.

CC InterPro; IPR000742; EGF\_2.

CC InterPro; IPR001881; EGF\_Ca.

CC InterPro; IPR001438; EGF\_II.

CC InterPro; IPR006209; EGF\_like.

CC InterPro; IPR002049; Laminin\_EGF.

CC InterPro; IPR008297; Notch.

CC InterPro; IPR000800; Notch\_dom.

CC Pfam; PF00023; ank; 6.

CC Pfam; PF00008; EGF; 35.

CC Pfam; PF00066; notch; 2.

CC PIRSF; PIRSF002279; Notch; 1.

CC PRINTS; PR00010; EGFLOOD.

CC PRINTS; PR00011; EGFLAMININ.

CC PRINTS; PR01452; NOTCH.

CC SMART; SM00248; ANK; 6.

CC SMART; SM00179; EGF\_CA; 23.

CC SMART; SM00004; NL; 2.

CC PROSITE; PS50297; ANK\_REPEAT; 1.

CC PROSITE; PS50088; ANK\_REPEAT; 4.

CC PROSITE; PS00010; ASX\_HYDROXYL; 22.

CC PROSITE; PS00022; EGF\_1; 34.

CC PROSITE; PS01186; EGF\_2; 29.

CC PROSITE; PS50026; EGF\_3; 35.

CC PROSITE; PS01187; EGF\_CA; 22.

CC Receptor; Transcription regulation; Activator; Differentiation; Developmental protein; Repeat; ANK repeat; EGF-like domain; Transmembrane; Glycoprotein; Signal; Phosphorylation.

CC SIGNAL 1 25 POTENTIAL.

CC CHAIN 26 2471 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 2.

CC CHAIN 1666 2471 NOTCH EXTRACELLULAR TRUNCATION (BY SIMILARITY).

CC CHAIN 1697 2471 NOTCH INTRACELLULAR DOMAIN (BY SIMILARITY).

CC DOMAIN 26 1677 EXTRACELLULAR (POTENTIAL).

CC TRANSMEM 1678 1698 POTENTIAL.

CC DOMAIN 1699 2471 CYTOPLASMIC (POTENTIAL).

CC DOMAIN 26 63 EGF-LIKE 1.

CC DOMAIN 64 102 EGF-LIKE 2.

CC DOMAIN 105 143 EGF-LIKE 3.

CC DOMAIN 144 180 EGF-LIKE 4.

CC DOMAIN 182 219 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).

FT	DOMAIN	221	258	EGF-LIKE 6.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	260	296	EGF-LIKE 7.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	298	336	EGF-LIKE 8.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	338	374	EGF-LIKE 9.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	375	413	EGF-LIKE 10.	
FT	DOMAIN	415	454	EGF-LIKE 11.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	456	492	EGF-LIKE 12.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	494	530	EGF-LIKE 13.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	532	568	EGF-LIKE 14.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	570	605	EGF-LIKE 15.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	607	643	EGF-LIKE 16.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	645	680	EGF-LIKE 17.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	682	718	EGF-LIKE 18.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	720	755	EGF-LIKE 19.	
FT	DOMAIN	757	793	EGF-LIKE 20.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	795	831	EGF-LIKE 21.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	833	871	EGF-LIKE 22.	
FT	DOMAIN	873	909	EGF-LIKE 23.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	911	947	EGF-LIKE 24.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	949	985	EGF-LIKE 25.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	987	1023	EGF-LIKE 26.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1025	1061	EGF-LIKE 27.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1063	1099	EGF-LIKE 28.	
FT	DOMAIN	1101	1147	EGF-LIKE 29.	
FT	DOMAIN	1149	1185	EGF-LIKE 30.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1187	1223	EGF-LIKE 31.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1225	1262	EGF-LIKE 32.	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1264	1302	EGF-LIKE 33.	
FT	DOMAIN	1304	1343	EGF-LIKE 34.	
FT	DOMAIN	1374	1412	EGF-LIKE 35.	
FT	REPEAT	1420	1456	LIN/NOTCH 1.	
FT	REPEAT	1503	1535	LIN/NOTCH 2.	
FT	REPEAT	1827	1871	ANK 1.	
FT	REPEAT	1876	1905	ANK 2.	
FT	REPEAT	1909	1939	ANK 3.	
FT	REPEAT	1943	1972	ANK 4.	
FT	REPEAT	1976	2005	ANK 5.	
FT	REPEAT	2009	2038	ANK 6.	
FT	DOMAIN	1645	1648	POLY-ALA.	
FT	DOMAIN	1994	1997	POLY-LEU.	
FT	DOMAIN	2426	2429	POLY-SER.	
FT	DISULFID	28	41	BY SIMILARITY.	
FT	DISULFID	35	51	BY SIMILARITY.	
FT	DISULFID	53	62	BY SIMILARITY.	
FT	DISULFID	68	79	BY SIMILARITY.	
FT	DISULFID	73	90	BY SIMILARITY.	
FT	DISULFID	92	101	BY SIMILARITY.	
FT	DISULFID	109	121	BY SIMILARITY.	
FT	DISULFID	115	131	BY SIMILARITY.	
FT	DISULFID	133	142	BY SIMILARITY.	
FT	DISULFID	148	159	BY SIMILARITY.	
FT	DISULFID	153	168	BY SIMILARITY.	
FT	DISULFID	170	179	BY SIMILARITY.	
FT	DISULFID	186	198	BY SIMILARITY.	
FT	DISULFID	192	207	BY SIMILARITY.	
FT	DISULFID	209	218	BY SIMILARITY.	
FT	DISULFID	225	236	BY SIMILARITY.	
FT	DISULFID	230	246	BY SIMILARITY.	
FT	DISULFID	248	257	BY SIMILARITY.	
Query Match 3.0%; Score 440; DB 1; Length 2471;					
Best Local Similarity 30.1%; Pred. No. 6.8e-17;					
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20					
QY	562	SVDNCPSN-CYNGDC---	ISG-TCHCFLGLGPDCCGRASCP-----	VLCSGNGQY	607
Db	795	NIDECASNPCLNQGTCTFDDISGYTCHCVLPYTGKNCQTVLAPCSPNPCENAAVCKESP	NF		854
QY	608	MKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTG--ITGT--CICNPGYKGESCEE--VD			661
Db	855	ESYTCLCAPGWQGRCTDIDECISKPCNMHGLCHNTQGSYMCECPGPGSGMDCEEDIDD			914
QY	662	CMDFTCSGRGVCVRG-----ECHCFVGWGGTNCETPRATCLDQ-CSGHGTF--LPDTGLCS			714

Db	915	CLANPCQNGGSCMDGVNTFTSCLCLPGFTGDKCQTMNECLSEPCKNGGTCSYVNSYTK	974
QY	715	CDPSWTGHDC--SIEICA-ADCGGHGVCVGG-----TCRCEDGWMGAAC---DQACHPR	763
Db	975	COAGFDGVHCENNINECTESSCFNGGTCTVDGINSFSLCPVGFTGFSCLHEINECSHP-	1033
QY	764	CAEHGTCRDG---KCECSPGWNGEHC-TIAHYLDRVVKEGCPGLCNGNGRCITLDLNGWH	818
Db	1034	CLNEGTCVDGLGTYRCSPLGYTGKNCQTLVNLCSR-----SPCKNKGTVCVQKKAESQ	1086
QY	819	CVCQLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPCCLQPLCHINPLCLGSPN	874
Db	1087	CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLCQHSVGCINAGN	1131
RESULT 6			
NOTC_DROME STANDARD; PRT; 2703 AA.			
AC	P07207	O97458; P04154; Q9W4T8;	
DT	01-NOV-1986	(Rel. 03, Created)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	15-MAR-2004	(Rel. 43, Last annotation update)	
DE	Neurogenic locus Notch protein precursor.		
GN	N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephyrroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Oregon-R; TISSUE=Embryo;		
RX	MEDLINE=86079539; PubMed=3935325;		
RA	Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;		
RT	"Nucleotide sequence from the neurogenic locus notch implies a gene		
RT	product that shares homology with proteins containing EGF-like		
RT	repeats.";		
RL	Cell 43:567-581(1985).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Canton-S, and Oregon-R; TISSUE=Embryo;		
RX	MEDLINE=87064624; PubMed=3097517;		
RA	Kidd S., Kelley M.R., Young M.W.;		
RT	"Sequence of the notch locus of Drosophila melanogaster: relationship		
RT	of the encoded protein to mammalian clotting and growth factors.";		
RL	Mol. Cell. Biol. 6:3094-3108(1986).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Berkeley;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,		
RA	Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,		

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R;  
RX MEDLINE=20196011; PubMed=10731137;  
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,  
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demalles J., Cadieu B.,  
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,  
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,  
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,  
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikioti F.,  
RA Beinert N., Dowe G., Schaefer U., Jaecckle H., Bucheton A.,  
RA Callister D.J., Campbell L.A., Darlamitsou A., Henderson N.S.,  
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,  
RA Glover D.M.;  
RT "From sequence to chromosome: the tip of the X chromosome of *D.*  
RT *melanogaster*,"  
RL Science 287:2220-2222(2000).  
RN [5]  
RP SEQUENCE OF 2505-2611 FROM N.A.  
RX MEDLINE=85099329; PubMed=2981631;  
RA Wharton K.A., Yedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;  
RT "opa: a novel family of transcribed repeats shared by the Notch locus  
RT and other developmentally regulated loci in *D. melanogaster*,"  
RL Cell 40:55-62(1985).  
RN [6]  
RP SEQUENCE OF 1-8 FROM N.A.  
RX MEDLINE=87257846; PubMed=3037327;  
RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;  
RT "Restriction of P-element insertions at the Notch locus of *Drosophila*  
RT *melanogaster*,"  
RL Mol. Cell. Biol. 7:1545-1548(1987).  
RN [7]  
RP INTERACTION WITH DX, AND MUTANT SU42C.  
RX MEDLINE=94215489; PubMed=8162848;  
RA Diederich R.J., Matsuno K., Hing H., Artavanis-Tsakonas S.;  
RT "Cytosolic interaction between *deltex* and Notch ankyrin repeats  
RT implicates *deltex* in the Notch signaling pathway,"  
RL Development 120:473-481(1994).  
RN [8]  
RP INTERACTION WITH DX.  
RX MEDLINE=95401878; PubMed=7671825;  
RA Matsuno K., Diederich R.J., Go M.J., Blaumueller C.M.,  
RA Artavanis-Tsakonas S.;  
RT "*Deltex* acts as a positive regulator of Notch signaling through  
RT interactions with the Notch ankyrin repeats,"  
RL Development 121:2633-2644(1995).  
RN [9]  
RP S3 CLEAVAGE BY PSN.  
RX MEDLINE=99221487; PubMed=10206646;  
RA Struhl G., Greenwald I.;  
RT "Presenilin is required for activity and nuclear access of Notch in  
RT *Drosophila*,"  
RL Nature 398:522-525(1999).  
RN [10]  
RP S3 CLEAVAGE BY PSN.  
RX MEDLINE=99221488; PubMed=10206647;  
RA Ye Y., Lukinova N., Fortini M.E.;

RT "Neurogenic phenotypes and altered Notch processing in *Drosophila*  
RT *Presenilin* mutants,"  
RL Nature 398:525-529(1999).  
RN [11]  
RP S2 CLEAVAGE BY KUZ.  
RX MEDLINE=21657146; PubMed=11799064;  
RA Lieber T., Kidd S., Young M.W.;  
RT "Kuzbanian-mediated cleavage of *Drosophila* Notch,"  
RL Genes Dev. 16:209-221(2002).  
RN [12]  
RP MUTANT MCD5.  
RX MEDLINE=21575956; PubMed=11719214;  
RA Ramain P., Khechumian K., Seugnet L., Arbogast N., Ackermann C.,  
RA Heitzler P.;  
RT "Novel Notch alleles reveal a *Deltex*-dependent pathway repressing  
RT neural fate,"  
RL Curr. Biol. 11:1729-1738(2001).  
RN [13]  
RP REVIEW.  
RX MEDLINE=22256570; PubMed=12369105;  
RA Portin P.;  
RT "General outlines of the molecular genetics of the Notch signalling  
RT pathway in *Drosophila melanogaster*: a review,"  
RL Hereditas 136:89-96(2002).  
CC -!- FUNCTION: Signaling protein, which regulates, with both positive  
CC and negative signals, the differentiation of at least central and  
CC peripheral nervous system and eye, wing disk, oogenesis, segmental  
CC appendages such as antennae and legs, and muscles, through lateral  
CC inhibition or induction. Functions as a receptor for membrane-  
CC bound ligands *Delta* and *Serrate* to regulate cell-fate  
CC determination. Upon ligand activation, and releasing from the cell  
CC membrane, the Notch intracellular domain (NICD) forms a  
CC transcriptional activator complex with Su(H) (Suppressor of  
CC hairless) and activates genes of the E(spl) complex. Essential for  
CC proper differentiation of ectoderm.  
CC -!- SUBUNIT: Interacts with Su(H) when activated. Interacts with Dx  
CC via its ANK repeats.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Upon activation and  
CC S3 cleavage, it is released from the cell membrane and enters into  
CC the nucleus in conjunction with Su(H).  
CC -!- PTM: Upon binding its ligands such as *Delta* or *Serrate*, it is  
CC cleaved (S2 cleavage) in its extracellular domain, close to the  
CC transmembrane domain. S2 cleavage is probably mediated by Kuz. It  
CC is then cleaved (S3 cleavage) downstream of its transmembrane  
CC domain, releasing it from the cell membrane. S3 cleavage requires  
CC Psn.  
CC -!- SIMILARITY: Belongs to the NOTCH family.  
CC -!- SIMILARITY: Contains 36 EGF-like domains.  
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.  
CC -!- SIMILARITY: Contains 6 ANK repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M16152; AAB59220.1; -.  
CC EMBL; M16153; AAB59220.1; JOINED.  
CC EMBL; M16149; AAB59220.1; JOINED.  
CC EMBL; M16150; AAB59220.1; JOINED.  
CC EMBL; M16151; AAB59220.1; JOINED.  
CC EMBL; K03508; AAA28725.1; -.  
CC EMBL; M13689; AAA28725.1; JOINED.  
CC EMBL; K03507; AAA28725.1; JOINED.  
CC EMBL; AE003426; AAF45848.2; -.  
CC EMBL; AL035436; CAB37610.1; -.  
CC EMBL; M12175; AAA74496.1; -.  
CC EMBL; M16025; AAA28726.1; -.

```
Query Match          2.9%; Score 438.5; DB 1; Length 2703;
Best Local Similarity 29.7%; Pred. No. 9.6e-17;
Matches 125; Conservative 50; Mismatches 165; Indels 81; Gaps 24;

QY 561 ESDVNCP--SNCYNGDCI-----SGTCHCFGLGFLGPDCCR-----ASCPVLCSGNGQYMK 609
Db 448 EDIDECDOGSPCEHNGICVNTPGSYRCNCSTGFTGPRCETNINECHSHP--CQNEGSCLD 505
QY 610 G-----RCLCHSGWKGAECDDVPTNQCIDVACSNHGTC---ITG-TCICNPGYKGESCE--E 659
Db 506 DPGTFRVCVMPGFTGTQCEIDIDEQQSNPCLNDGTCHDKINGFKCSALGFTGARQQINI 565
QY 660 VDCMDPTCSGRGVC---VRG-ECHCFVGGWGTNCETPRATCLDQCSGHGTFPLDTG--LC 713
Db 566 DDCQSOPCRNRGICHDSIAGYSCECPPGYTGTSCEININDCDNSPCHRGKCIDDDVNSFKC 625
QY 714 SCDPSWTGHDCSIEI--CAAD-CGGHGVG---VGG-TCRCEDGWMGAACDQRA--CHPR- 763
Db 626 LCDPGYTYICQKQINECESNPCQFDGHQCDRVGSYYCQCQAGTSGKNCEVNVNECHSNP 685
QY 764 CAEHGTCRDG----KCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNRCRTLDLNGWHC 819
Db 686 CUNGATCIDGINSYKQCQVPGFTGQHC-----EKNVDECISSPCANNGVCIDQVNGYKC 739
QY 820 VCQLGWRGAGCDTSMETACGDSKNDG---DGLVDCMDPDCCQLPLCHINPLCLGSPNPL 876
Db 740 ECPRGFYDAHCLSDVDECASNPCVNEGRCEDEGINF-----ICHCPPGYTGKRCEL 790
QY 877 DIIQETQVPVSQQNLHSPFYDRIKFLVGRDSTHPIPG-----ENPFDGGHAC 922
Db 791 DIDECSNPPC--QHGGTCYDKLNAF---SCQCMPGYTGKCEITNIDDCVTNPGCGNGGTC 844
QY 923 V 923
Db 845 I 845

RESULT 7
JAG3_BRARE          STANDARD; PRT; 1213 AA.
ID JAG3 BRARE
AC Q90Y54;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Jagged 3 precursor (Jagged3).
GN JAG3.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Oda T., Chandrasekharappa S.C.;
RT "Isolation, characterization and expression analysis of Zebrafish
RT Jagged genes.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Ligand for Notch receptors and involved in the mediation
CC of Notch signaling (By similarity). Seems to be involved in cell-
CC fate decisions.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- SIMILARITY: Contains 17 EGF-like domains.
CC -!- SIMILARITY: Contains 1 DSL domain.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AF229451; AAL08216.1; -.
DR ZFIN; ZDB-GENE-011128-4; Jag3.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR001774; DSL.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR009041; PMP_inhibitor.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF01414; DSL; 1.
DR Pfam; PF00008; EGF; 14.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_CA; 10.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 10.
DR PROSITE; PS00022; EGF_1; 16.
DR PROSITE; PS01186; EGF_2; 12.
DR PROSITE; PS00026; EGF_3; 15.
DR PROSITE; PS01187; EGF_CA; 8.
DR PROSITE; PS01208; VWFC_1; FALSE NEG.
KW Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
KW Repeat; Transmembrane; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1213 JAGGED 3.
FT DOMAIN 27 1064 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1065 1087 POTENTIAL.
FT DOMAIN 1088 1213 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 164 226 DSL.
FT DOMAIN 227 260 EGF-LIKE 1.
FT DOMAIN 258 291 EGF-LIKE 2.
FT DOMAIN 293 331 EGF-LIKE 3.
FT DOMAIN 333 369 EGF-LIKE 4.
FT DOMAIN 371 407 EGF-LIKE 5.
FT DOMAIN 409 445 EGF-LIKE 6.
FT DOMAIN 447 482 EGF-LIKE 7.
FT DOMAIN 484 520 EGF-LIKE 8.
FT DOMAIN 522 558 EGF-LIKE 9.
FT DOMAIN 592 624 EGF-LIKE 10.
FT DOMAIN 626 662 EGF-LIKE 11.
FT DOMAIN 664 700 EGF-LIKE 12.
FT DOMAIN 702 738 EGF-LIKE 13.
FT DOMAIN 746 777 EGF-LIKE 14.
FT DOMAIN 779 815 EGF-LIKE 15.
FT DOMAIN 817 853 EGF-LIKE 16.
FT DOMAIN 860 914 VWFC.
FT DOMAIN 918 956 EGF-LIKE 17.
FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 235 248 BY SIMILARITY.
FT DISULFID 250 259 BY SIMILARITY.
FT DISULFID 262 273 BY SIMILARITY.
FT DISULFID 268 279 BY SIMILARITY.
FT DISULFID 281 290 BY SIMILARITY.
FT DISULFID 297 309 BY SIMILARITY.
FT DISULFID 303 319 BY SIMILARITY.
FT DISULFID 321 330 BY SIMILARITY.
FT DISULFID 337 348 BY SIMILARITY.
FT DISULFID 342 357 BY SIMILARITY.
FT DISULFID 359 368 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 395 BY SIMILARITY.
FT DISULFID 397 406 BY SIMILARITY.
FT DISULFID 413 424 BY SIMILARITY.
FT DISULFID 418 433 BY SIMILARITY.
FT DISULFID 435 444 BY SIMILARITY.
FT DISULFID 451 461 BY SIMILARITY.
FT DISULFID 455 470 BY SIMILARITY.
FT DISULFID 472 481 BY SIMILARITY.
FT DISULFID 488 499 BY SIMILARITY.
FT DISULFID 493 508 BY SIMILARITY.
```

FT	DISULFID	510	519	BY SIMILARITY.
FT	DISULFID	526	537	BY SIMILARITY.
FT	DISULFID	531	546	BY SIMILARITY.
FT	DISULFID	548	557	BY SIMILARITY.
FT	DISULFID	641	630	BY SIMILARITY.
FT	DISULFID	635	650	BY SIMILARITY.
FT	DISULFID	661	652	BY SIMILARITY.
FT	DISULFID	668	679	BY SIMILARITY.
FT	DISULFID	673	688	BY SIMILARITY.
FT	DISULFID	690	699	BY SIMILARITY.
FT	DISULFID	706	717	BY SIMILARITY.
FT	DISULFID	711	726	BY SIMILARITY.
FT	DISULFID	728	737	BY SIMILARITY.
FT	DISULFID	783	794	BY SIMILARITY.
FT	DISULFID	788	803	BY SIMILARITY.
FT	DISULFID	805	814	BY SIMILARITY.
FT	DISULFID	821	832	BY SIMILARITY.
FT	DISULFID	826	841	BY SIMILARITY.
FT	DISULFID	843	852	BY SIMILARITY.
FT	DOMAIN	938	941	POLY-PRO.
FT	CARBOHYD	139	139	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	214	214	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	556	556	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	742	742	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	957	957	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	988	988	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1042	1042	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE	1213 AA; 133365 MW; 505F16A7E20D9534 CRC64;		
Query Match 2.9%; Score 434; DB 1; Length 1213;				
Best Local Similarity 27.2%; Pred. No. 4.8e-17;				
Matches 134; Conservative 48; Mismatches 177; Indels 134; Gaps 27;				
QY	496 LLDGRRLLTQEARSLGTPRQSRGTVP-----PSSHETGFIQYLD-----SG-----	537		
Db	98 VLGGNKFSYKGRS-----EKSRIVLFPFSAWPRSY-TLIVEALDFNETASESGKLIK 151			
QY	538 IWHLAFYNDKE-----SEVVSF-----LTTAIESV--DNCPSNC-----YNGDC-I 577			
Db	152 AYHSGMINENQWQRLTHNGPVAQFYQIRVTCLEHYHGGCNKFCRPRDEFHGYTCDQ 211			
QY	578 SGTCHCFLGLPDCGRASCPVLC-----GNGQVMKG-----	610		
Db	212 NGNKTCLGWTGPDNTAICRQGCSTEHGSCQKPGGCKLYGWQGPYCDKICPHGCVHG 271			
QY	611 -----RLCHSGWKGAECVPTNQC-IDVACSNHGTCTI-TG-----TCICNPYKGESCE 658			
Db	272 TCVEPWQCLCDTNWGGQLCDKLDLYCGTHQPCNLNGGTCNTGPDKYQCSCEDGYSGVNC 331			
QY	659 EVD--CMDPTCSGRGVCVRG-----BCHCFVWGVTNCTPRATCL-DQCSGHGTF--LPD 709			
Db	332 RAEHACLSNFCANGGTCKETSQGYECHCAIGWSGTSCEINVDCTPNQCKHGGTCQDLVN 391			
QY	710 TGLCSCDPSWTGHDCSIE-----ICAADCGHGVCGGTCRCEDGWMGAACDQRA- 759			
Db	392 GFKACAPPHWTGKTCQIDANECEKPCVNAKSC--HNLIGAYFCECLPGWSGQNCININ 449			
QY	760 -CHPRCAEHGTCRD---GKCECSPGWNGEHCTIAHYLDYRVVKEGCPGLCNGNGRCTLDL 814			
Db	450 DCKGQCLNGGTCCKDLVNGYRCCLPPGYTGEQC-----EKDVDECASSPCLNGRCQDEV 503			
QY	815 NGWHVCVQLGWRGAGCDTSMETACGDSKNDG-----DGLVDCMD-----PD 856			
Db	504 NGFQCLCPAGFSGLCQLDLID-YCKPNPCQNGAQCFNLASDYFCCKPDDYEGKNCSHLKD 562			
QY	857 CCLQPLCHINPLC 869			
Db	563 HCRTTSCQVIDSC 575			
RESULT 8				
NOTC_XENLA				
ID	NOTC_XENLA	STANDARD;	PRT;	2524 AA.

AC	P21783;			
DT	01-MAY-1991 (Rel. 18, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neurogenic locus notch protein homolog precursor (XOTCH protein).			
GN	XOTCH.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90385285; PubMed=2402639;			
RA	Coffman C., Harris W., Kintner C.;			
RT	"Xotch, the Xenopus homolog of Drosophila notch.";			
RL	Science 249:1438-1441(1990).			
RN	[2]			
RP	REVISIONS TO 1759-1782.			
RA	Kintner C.;			
RL	Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-!- DEVELOPMENTAL STAGE: Expressed almost uniformly in early embryos.			
CC	-!- SIMILARITY: Belongs to the NOTCH family.			
CC	-!- SIMILARITY: Contains 36 EGF-like domains.			
CC	-!- SIMILARITY: Contains 3 Lin/Notch repeats.			
CC	-!- SIMILARITY: Contains 6 ANK repeats.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M33874; AAB02039.1; -			
DR	HSSP; P00740; IEDM.			
DR	InterPro; IPR002110; ANK.			
DR	InterPro; IPR000152; Asx_hydroxyl_s.			
DR	InterPro; IPR000742; EGF_2.			
DR	InterPro; IPR001981; EGF_Ca.			
DR	InterPro; IPR001438; EGF_II.			
DR	InterPro; IPR006209; EGF_like.			
DR	InterPro; IPR002049; Laminin_EGF.			
DR	InterPro; IPR008297; Notch.			
DR	InterPro; IPR000800; Notch_dom.			
DR	Pfam; PF00023; ank; 6.			
DR	Pfam; PF00008; EGF; 36.			
DR	Pfam; PF00066; notch; 3.			
DR	Pfam; PIRSF002279; Notch; 1.			
DR	PRINTS; PR00010; EGF_BLOOD.			
DR	PRINTS; PR00011; EGFLAMININ.			
DR	PRINTS; PR01452; NOTCH.			
DR	SMART; SM00248; ANK; 6.			
DR	SMART; SM00179; EGF_CA; 24.			
DR	SMART; SM00004; NL; 2.			
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.			
DR	PROSITE; PS50088; ANK_REPEAT; 4.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 23.			
DR	PROSITE; PS00022; EGF_1; 34.			
DR	PROSITE; PS01186; EGF_2; 29.			
DR	PROSITE; PS50026; EGF_3; 36.			
DR	PROSITE; PS01187; EGF_CA; 21.			
KW	Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;			
KW	Transmembrane; Signal; Glycoprotein.			
FT	SIGNAL 1 19			POTENTIAL.
FT	CHAIN 20 2524			NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.
FT	DOMAIN 20 1728			EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1729 1750			POTENTIAL.
FT	DOMAIN 1751 2524			CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 20 57			EGF-LIKE 1.
FT	DOMAIN 58 99			EGF-LIKE 2.

FT	DOMAIN	102	140	EGF-LIKE 3.		FT	DISULFID	415	428	BY SIMILARITY.
FT	DOMAIN	141	177	EGF-LIKE 4.		FT	DISULFID	422	437	BY SIMILARITY.
FT	DOMAIN	179	215	EGF-LIKE 5.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	439	448	BY SIMILARITY.
FT	DOMAIN	217	254	EGF-LIKE 6.		FT	DISULFID	455	466	BY SIMILARITY.
FT	DOMAIN	256	292	EGF-LIKE 7.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	460	475	BY SIMILARITY.
FT	DOMAIN	294	332	EGF-LIKE 8.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	477	486	BY SIMILARITY.
FT	DOMAIN	334	370	EGF-LIKE 9.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	493	504	BY SIMILARITY.
FT	DOMAIN	371	409	EGF-LIKE 10.		FT	DISULFID	498	513	BY SIMILARITY.
FT	DOMAIN	411	449	EGF-LIKE 11.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	515	524	BY SIMILARITY.
FT	DOMAIN	451	487	EGF-LIKE 12.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	531	542	BY SIMILARITY.
FT	DOMAIN	489	525	EGF-LIKE 13.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	536	551	BY SIMILARITY.
FT	DOMAIN	527	563	EGF-LIKE 14.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	553	562	BY SIMILARITY.
FT	DOMAIN	565	600	EGF-LIKE 15.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	569	579	BY SIMILARITY.
FT	DOMAIN	602	638	EGF-LIKE 16.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	574	588	BY SIMILARITY.
FT	DOMAIN	640	675	EGF-LIKE 17.		FT	DISULFID	590	599	BY SIMILARITY.
FT	DOMAIN	677	713	EGF-LIKE 18.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	606	617	BY SIMILARITY.
FT	DOMAIN	715	750	EGF-LIKE 19.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	611	626	BY SIMILARITY.
FT	DOMAIN	752	788	EGF-LIKE 20.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	628	637	BY SIMILARITY.
FT	DOMAIN	790	826	EGF-LIKE 21.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	644	654	BY SIMILARITY.
FT	DOMAIN	828	866	EGF-LIKE 22.		FT	DISULFID	649	663	BY SIMILARITY.
FT	DOMAIN	868	904	EGF-LIKE 23.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	665	674	BY SIMILARITY.
FT	DOMAIN	906	942	EGF-LIKE 24.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	681	692	BY SIMILARITY.
FT	DOMAIN	944	980	EGF-LIKE 25.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	686	701	BY SIMILARITY.
FT	DOMAIN	982	1018	EGF-LIKE 26.		FT	DISULFID	703	712	BY SIMILARITY.
FT	DOMAIN	1020	1056	EGF-LIKE 27.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	719	729	BY SIMILARITY.
FT	DOMAIN	1058	1094	EGF-LIKE 28.		FT	DISULFID	724	738	BY SIMILARITY.
FT	DOMAIN	1096	1142	EGF-LIKE 29.		FT	DISULFID	740	749	BY SIMILARITY.
FT	DOMAIN	1144	1180	EGF-LIKE 30.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	756	767	BY SIMILARITY.
FT	DOMAIN	1182	1218	EGF-LIKE 31.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	761	776	BY SIMILARITY.
FT	DOMAIN	1220	1264	EGF-LIKE 32.	CALCIUM-BINDING (POTENTIAL).	FT	DISULFID	778	787	BY SIMILARITY.
FT	DOMAIN	1266	1304	EGF-LIKE 33.		FT	DISULFID	794	805	BY SIMILARITY.
FT	DOMAIN	1306	1346	EGF-LIKE 34.		FT	DISULFID	799	814	BY SIMILARITY.
FT	DOMAIN	1347	1383	EGF-LIKE 35.		FT	DISULFID	816	825	BY SIMILARITY.
FT	DOMAIN	1386	1424	EGF-LIKE 36.		FT	DISULFID	832	843	BY SIMILARITY.
FT	REPEAT	1441	1478	LIN/NOTCH 1.		FT	DISULFID	837	854	BY SIMILARITY.
FT	REPEAT	1479	1520	LIN/NOTCH 2.		FT	DISULFID	856	865	BY SIMILARITY.
FT	REPEAT	1521	1560	LIN/NOTCH 3.		FT	DISULFID	872	883	BY SIMILARITY.
FT	REPEAT	1876	1919	ANK 1.		FT	DISULFID	877	892	BY SIMILARITY.
FT	REPEAT	1924	1953	ANK 2.		FT	DISULFID	894	903	BY SIMILARITY.
FT	REPEAT	1957	1987	ANK 3.		FT	DISULFID	910	921	BY SIMILARITY.
FT	REPEAT	1991	2020	ANK 4.		FT	DISULFID	915	930	BY SIMILARITY.
FT	REPEAT	2024	2053	ANK 5.		FT	DISULFID	932	941	BY SIMILARITY.
FT	REPEAT	2057	2086	ANK 6.		FT	DISULFID	986	997	BY SIMILARITY.
FT	DISULFID	22	35	BY SIMILARITY.		FT	DISULFID	991	1006	BY SIMILARITY.
FT	DISULFID	29	45	BY SIMILARITY.		FT	DISULFID	1008	1017	BY SIMILARITY.
FT	DISULFID	47	56	BY SIMILARITY.		FT	DISULFID	1024	1035	BY SIMILARITY.
FT	DISULFID	62	74	BY SIMILARITY.		FT	DISULFID	1029	1044	BY SIMILARITY.
FT	DISULFID	68	87	BY SIMILARITY.		FT	DISULFID	1046	1055	BY SIMILARITY.
FT	DISULFID	89	98	BY SIMILARITY.		FT	DISULFID	1062	1073	BY SIMILARITY.
FT	DISULFID	106	117	BY SIMILARITY.		FT	DISULFID	1067	1082	BY SIMILARITY.
FT	DISULFID	111	128	BY SIMILARITY.		FT	DISULFID	1084	1093	BY SIMILARITY.
FT	DISULFID	130	139	BY SIMILARITY.		FT	DISULFID	1084	1093	BY SIMILARITY.
FT	DISULFID	145	156	BY SIMILARITY.		FT	DISULFID	1100	1121	BY SIMILARITY.
FT	DISULFID	150	165	BY SIMILARITY.		FT	DISULFID	1100	1121	BY SIMILARITY.
FT	DISULFID	167	176	BY SIMILARITY.		FT	DISULFID	1115	1130	BY SIMILARITY.
FT	DISULFID	183	194	BY SIMILARITY.						
FT	DISULFID	188	203	BY SIMILARITY.						
FT	DISULFID	205	214	BY SIMILARITY.						
FT	DISULFID	221	232	BY SIMILARITY.						
FT	DISULFID	226	242	BY SIMILARITY.						
FT	DISULFID	244	253	BY SIMILARITY.						
FT	DISULFID	260	271	BY SIMILARITY.						
FT	DISULFID	265	280	BY SIMILARITY.						
FT	DISULFID	282	291	BY SIMILARITY.						
FT	DISULFID	298	311	BY SIMILARITY.						
FT	DISULFID	305	320	BY SIMILARITY.						
FT	DISULFID	322	331	BY SIMILARITY.						
FT	DISULFID	338	349	BY SIMILARITY.						
FT	DISULFID	343	358	BY SIMILARITY.						
FT	DISULFID	360	369	BY SIMILARITY.						
FT	DISULFID	375	386	BY SIMILARITY.						
FT	DISULFID	380	397	BY SIMILARITY.						
FT	DISULFID	399	408	BY SIMILARITY.						

Query Match

2.9%; Score 429; DB 1; Length 2524;

Best Local Similarity 30.6%; Pred. No. 3.1e-16;

Matches 112; Conservative 38; Mismatches 144; Indels 72; Gaps 21;

QY	562	SVDNCPN-CYNGNDCISG----	---TCHCFLGFLGPDCCR-----	ASCPVLCSGNGQYMKG-	610	
Db	489	NIDECASNPCLNHGKCIDKINEFRCD	CPTGFSNLCQHDDECTSTP--	CKNGAKCLDGP	546	
QY	611	---RCLCHSGWKGAECDVPTNQCIDV	ACSNHGTCITG----	TCICNPGYKGESCEE--	VD 661	
Db	547	NSYTCQCTEGFTGRHCEQDINECI	PDPC-HYGTCKDGIATFTCLCRPGYT	TRLCDNDINE	605	
QY	662	CMDPTCSGRGVCVRGE---	CHCFVGGWGTNCET-----	PRATCLDQCSGHGTF	707	
Db	606	CLSKPCLNGGQCTDRENGYICTCP	KGTGTGVNCE	TXIDKIDKIDGVE---	662	
QY	708	PDTGLCSDPSPSWTGHDCSIEICAAD	---	CGGHGVCV----	GGTCRCEDGWMGAAC--	DQR 758
Db	663	-----CTCEPGYTGKLCNININE	CDNPNCRNGG	TKDQINGFT	CVCPDGYHDMCLSEVN	717





RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;  
RX MEDLINE=931178563; PubMed=8440332;  
RA Lardelli M., Lendahl U.;  
RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a  
RL wide variety of tissues";  
RN Exp. Cell Res. 204:364-372(1993).  
[5]  
RP SEQUENCE OF 1659-1673 FROM N.A.  
RX MEDLINE=99364499; PubMed=10437788;  
RA Lee J.S., Ishimoto A., Yanagawa S.I.;  
RT "Murine leukemia provirus-mediated activation of the Notch1 gene leads  
RL to induction of HES-1 in a mouse T lymphoma cell line, DL-3.";  
RN FEBS Lett. 455:276-280(1999).  
[6]  
RP SEQUENCE OF 1950-2201 FROM N.A.  
RX MEDLINE=98029496; PubMed=9384671;  
RA Messerle M., Follo M., Nehls M., Eggert H., Boehm T.;  
RT "Dynamic changes in gene expression during in vitro differentiation of  
RL mouse embryonic stem cells.";  
RN Cytokines Cell. Mol. Ther. 1:139-143(1995).  
[7]  
RP SEQUENCE OF 1655-1659. CLEAVAGE BY FURIN-LIKE CONVERTASE, AND  
RX MUTAGENESIS OF 1651-ARG--ARG-1654.  
RA Logeat F., Bessia C., Brou C., LeBail O., Jarriault S., Seidah N.G.,  
RT Israel A.;  
RL "The Notch1 receptor is cleaved constitutively by a furin-like  
RN convertase.";  
RP Proc. Natl. Acad. Sci. U.S.A. 95:8108-8112(1998).  
[8]  
RP PARTIAL SEQUENCE, AND POST-TRANSLATIONAL PROCESSING.  
RX MEDLINE=21523956; PubMed=11518718;  
RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;  
RT "Murine notch homologs (NL-4) undergo presenilin-dependent  
RL proteolysis.";  
RN J. Biol. Chem. 276:40268-40273(2001).  
[9]  
RP POST-TRANSLATIONAL PROCESSING.  
RX MEDLINE=21374376; PubMed=11459941;  
RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;  
RT "Conservation of the biochemical mechanisms of signal transduction  
RL among mammalian Notch family members.";  
RN Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).  
[10]  
RP INTERACTION WITH DTX1 AND DTX2.  
RX MEDLINE=21123790; PubMed=11226752;  
RA Kishi N., Tang Z., Maeda Y., Hirai A., Mo R., Ito M., Suzuki S.,  
RA Nakao K., Kinoshita T., Kadesch T., Hui C.-C., Artavanis-Tsakonas S.,  
RA Okano H., Matsuno K.;  
RT "Murine homologs of deltex define a novel gene family involved in  
RL vertebrate Notch signaling and neurogenesis.";  
RN Int. J. Dev. Neurosci. 19:21-35(2001).  
CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands  
CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.  
CC Upon ligand activation through the released notch intracellular  
CC domain (NICD) it forms a transcriptional activator complex with  
CC RBP-J kappa and activates genes of the enhancer of split locus.  
CC Affects the implementation of differentiation, proliferation and  
CC apoptotic programs (By similarity). May play an essential role in  
CC postimplantation development, probably in some aspect of cell  
CC specification and/or differentiation. May be involved in mesoderm  
CC development, somite formation and neurogenesis. Involved in the  
CC maturation of both CD4+ and CD8+ cells in the thymus.  
CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-  
CC terminal fragment N(EC) which are probably linked by disulfide  
CC bonds. Interacts with DTX1 and DTX2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following  
CC proteolytical processing NICD is translocated to the nucleus.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q01705-1; Sequence=Displayed;  
CC Name=2;

CC IsoId=Q01705-2; Sequence=VSP\_001402, VSP\_001403, VSP\_001404;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Highly expressed in the brain, lung and  
CC thymus. Expressed at lower levels in the spleen, bone-marrow,  
CC spinal cord, eyes, mammary gland, liver, intestine, skeletal  
CC muscle, kidney and heart.  
CC -!- DEVELOPMENTAL STAGE: First detected in the mesoderm at 7.5 dpc By  
CC 8.5 dpc highly expressed in presomitic mesoderm, mesenchyme and  
CC endothelial cells, while much lower levels are seen in the  
CC neuroepithelium. Between 9.5-10.5 dpc expressed at high levels in  
CC the neuroepithelium. At 13.5 dpc expressed in the surface  
CC ectoderm, eye and developing whisker follicles.  
CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form  
CC which is proteolytically cleaved by a furin-like convertase in the  
CC trans-Golgi network before it reaches the plasma membrane to yield  
CC an active, ligand-accessible form. Cleavage results in a C-  
CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following  
CC ligand binding, it is cleaved by TNF-alpha converting enzyme  
CC (TACE) to yield a membrane-associated intermediate fragment called  
CC notch extracellular truncation (NEXT). This fragment is then  
CC cleaved by presenilin dependent gamma-secretase to release a  
CC notch-derived peptide containing the intracellular domain (NICD)  
CC from the membrane.  
CC -!- PTM: Phosphorylated.  
CC -!- SIMILARITY: Belongs to the NOTCH family.  
CC -!- SIMILARITY: Contains 36 EGF-like domains.  
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.  
CC -!- SIMILARITY: Contains 5 ANK repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z11886; CAA77941.1; -;  
DR EMBL; L02613; AAK14898.1; -;  
DR EMBL; X68278; CAA48339.1; -;  
DR EMBL; AJ238029; CAB40733.1; -;  
DR EMBL; X82562; CAA57909.1; -;  
DR PIR; A46019; A46019.  
DR PIR; B49175; B49175.  
DR HSSP; P00740; 1EDM.  
DR MGD; MGI:97363; Notch1.  
DR GO; GO:0005887; C:integral to plasma membrane; IC.  
DR GO; GO:0005515; P:protein binding; IPI.  
DR GO; GO:0030154; P:cell differentiation; IMP.  
DR GO; GO:0007386; P:compartment specification; IMP.  
DR GO; GO:0007219; P:N signaling pathway; IC.  
DR GO; GO:0045944; P:positive regulation of transcription from P. . .; IDA.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000152; Asx\_hydroxyl\_S.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR008297; Notch.  
DR InterPro; IPR000800; Notch\_dom.  
DR Pfam; PF00023; ank; 7.  
DR Pfam; PF00008; EGF; 35.  
DR Pfam; PF00066; notch; 3.  
DR PIRSF; PIRSF002279; Notch; 1.  
DR PRINTS; PR00010; EGFBLLOOD.  
DR PRINTS; PR00011; EGFLAMININ.  
DR PRINTS; PR01452; NOTCH.  
DR SMART; SM00248; ANK; 6.  
DR SMART; SM00179; EGF\_CA; 24.  
DR SMART; SM00004; NL; 2.  
DR PROSITE; PS50297; ANK\_REPEAT; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 2.



CC from the membrane.  
CC -!- PTM: Phosphorylated.  
CC -!- SIMILARITY: Belongs to the NOTCH family.  
CC -!- SIMILARITY: Contains 35 EGF-like domains.  
CC -!- SIMILARITY: Contains 2 Lin/Notch repeats.  
CC -!- SIMILARITY: Contains 6 ANK repeats.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; D32210; BAA22094.1; --  
DR EMBL; X68279; CAA48340.1; --  
DR EMBL; U31881; AAC52924.1; --  
DR PIR; A49175; A49175.  
DR HSSP; P16109; 1FSB.  
DR MGD; MGI:97364; Notch2.  
DR GO; GO:0005887; C:integral to plasma membrane; IC.  
DR GO; GO:0005515; F:protein binding; IPI.  
DR GO; GO:0002011; P:morphogenesis of an epithelial sheet; IMP.  
DR GO; GO:0007219; P:N signaling pathway; IC.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000152; Asx hydroxyl S.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.  
DR InterPro; IPR006209; EGF\_Like.  
DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR008297; Notch.  
DR InterPro; IPR000800; Notch\_dom.  
DR Pfam; PF00023; ank; 6.  
DR Pfam; PF00008; EGF; 34.  
DR Pfam; PF00066; notch; 2.  
DR PIRSF; PIRSF002279; Notch; 1.  
DR PRINTS; PR00010; EGFLOOD.  
DR PRINTS; PR00011; EGF\_LAMININ.  
DR PRINTS; PR01452; NOTCH.  
DR SMART; SM00248; ANK; 6.  
DR SMART; SM00179; EGF\_CA; 23.  
DR SMART; SM00004; NL; 3.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 22.  
DR PROSITE; PS00022; EGF\_1; 33.  
DR PROSITE; PS01186; EGF\_2; 27.  
DR PROSITE; PS50026; EGF\_3; 35.  
DR PROSITE; PS01187; EGF\_CA; 22.  
KW Receptor; Transcription regulation; Activator; Differentiation;  
KW Developmental protein; Repeat; ANK repeat; EGF-like domain;  
KW Transmembrane; Glycoprotein; Signal; Phosphorylation;  
KW Alternative splicing.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 2470 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 2.  
FT CHAIN 1666 2470 NOTCH EXTRACELLULAR TRUNCATION.  
FT CHAIN 1697 2470 NOTCH INTRACELLULAR DOMAIN.  
FT DOMAIN 26 1677 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1678 1698 POTENTIAL.  
FT DOMAIN 1699 2470 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 26 63 EGF-LIKE 1.  
FT DOMAIN 64 102 EGF-LIKE 2.  
FT DOMAIN 105 143 EGF-LIKE 3.  
FT DOMAIN 144 180 EGF-LIKE 4.  
FT DOMAIN 182 219 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 221 256 EGF-LIKE 6 (INCOMPLETE).  
FT DOMAIN 258 294 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 296 334 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 336 372 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 373 411 EGF-LIKE 10.  
FT DOMAIN 413 452 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).

FT	DOMAIN	454	490	EGF-LIKE 12,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	492	528	EGF-LIKE 13,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	530	566	EGF-LIKE 14,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	568	603	EGF-LIKE 15,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	605	641	EGF-LIKE 16,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	643	678	EGF-LIKE 17,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	680	716	EGF-LIKE 18,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	718	753	EGF-LIKE 19,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	755	791	EGF-LIKE 20,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	793	829	EGF-LIKE 21,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	831	869	EGF-LIKE 22,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	871	907	EGF-LIKE 23,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	909	945	EGF-LIKE 24,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	947	983	EGF-LIKE 25,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	985	1021	EGF-LIKE 26,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1023	1059	EGF-LIKE 27,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1061	1097	EGF-LIKE 28,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1099	1145	EGF-LIKE 29,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1147	1183	EGF-LIKE 30,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1185	1221	EGF-LIKE 31,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1223	1260	EGF-LIKE 32,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1262	1300	EGF-LIKE 33,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1302	1345	EGF-LIKE 34,	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	1372	1410	EGF-LIKE 35,	CALCIUM-BINDING (POTENTIAL).
FT	REPEAT	1418	1454	LIN/NOTCH 1.	
FT	REPEAT	1501	1533	LIN/NOTCH 2.	
FT	REPEAT	1825	1869	ANK 1.	

Query Match 2.8%; Score 424; DB 1; Length 2470;  
Best Local Similarity 29.6%; Pred. No. 5.8e-16;  
Matches 105; Conservative 44; Mismatches 146; Indels 60; Gaps 19;  
  
QY 562 SVDNCPN-CYNGDC---ISG-TCHCFLGFLGPDGGRSCP-----VLCGNGQY 607  
Db 793 NIDECASNPLNQGTCTFDDVSGYTHCMLPYTKNCQTVLAPCSNPNCENAAVCKEAPNF 852  
  
QY 608 MKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTCI---TGTCICNPYKGESCEE--VD 661  
Db 853 ESFSCLCAPGWQGRKCTVDVDECIKSPCMNNGVCHNTQGSYVCECPFGSGMDCDEEDIND 912  
  
QY 662 CMDPTCSGRGVCV---RGECHCFVWGWTNCTETPRATCLDQ-CSGHGTF--LPDTGLCS 714  
Db 913 CLANPCQNGSGCVDHVNTFSCQCHPGFIDGKQCTDMNECLSEPCNKNGGTCSYVNSYTCT 972  
  
QY 715 CDPSTWTHDC--SIEICA-ADCGHGVVGVGG---TCRCEDGWMGAAC--DQRACHPR-C 764  
Db 973 CPAGFHGVHCENNIDECTESSCFNGGTCVDGINSFSCLCVPVGTGPFCLHDINECSSNPC 1032  
  
QY 765 AEHGTCTRDG---KCECSPGMNGEHC-TIAHYLDVVVKEGCPGLCNGNGRCTLDLNGWHC 819  
Db 1033 LNAGTCVDGLGTYRCICPLGYTGKNCQTLVNLCSR-----SPCKNKGTCTVQEKARPHC 1085  
  
QY 820 VCQLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPDCCCLQPLCHINPLCLGSPN 874  
Db 1086 LCPPGWDGAYCDV-LNVSKAA-----ALQKGVPEHLQHSQHCINAGN 1129

RESULT 12  
FBP1\_STRPU  
ID FBP1\_STRPU STANDARD; PRT; 1064 AA.  
AC P10079;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-FEB-1996 (Rel. 33, last sequence update)  
DT 15-MAR-2004 (Rel. 43, last annotation update)  
DE Fibropellin I precursor (Epidermal growth factor-related protein 1)  
DE (UEGF-1).  
GN EGF1.  
OS Strongylocentrotus purpuratus (Purple sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;  
OC Strongylocentrotus.  
OX NCBI\_TaxID=7668;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=90112459; PubMed=2514273;  
RA Delgadillo-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.;  
RT "Structural analysis of the uEGF gene in the sea urchin  
strongylocentrotus purpuratus reveals more similarity to vertebrate  
than to invertebrate genes with EGF-like repeats.";  
RL J. Mol. Evol. 29:314-327(1989).  
RN [2]  
RP SEQUENCE OF 279-476 AND 781-1064 FROM N.A.  
RX MEDLINE=87319677; PubMed=3498216;  
RA Hursh D.A., Andrews M.E., Raff R.A.;  
RT "A sea urchin gene encodes a polypeptide homologous to epidermal  
growth factor.";  
RL Science 237:1487-1490(1987).  
RN [3]  
RP AVIDIN-LIKE DOMAIN.  
RX MEDLINE=89196806; PubMed=2784773;  
RA Hunt L.T., Barker W.C.;  
RT "Avidin-like domain in an epidermal growth factor homolog from a sea  
urchin.";  
RL FASEB J. 3:1760-1764(1989).  
RN [4]  
RP CHARACTERIZATION.  
RX MEDLINE=91285254; PubMed=2060714;  
RA Bisgrove B.W., Andrews M.E., Raff R.A.;  
RT "Fibropellins, products of an EGF repeat-containing gene, form a  
unique extracellular matrix structure that surrounds the sea urchin  
embryo.";  
RL Dev. Biol. 146:89-99(1991).  
CC -!- FUNCTION: Form the apical lamina, a component of the extracellular  
matrix.  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR. IN VESICLES IN THE CYTOPLASM  
OF UNFERTILIZED EGGS, THEN TO THE BASE OF THE HYALIN LAYER  
THROUGHOUT DEVELOPMENT AND FINALLY IN THE APICAL LAMINA IN LATE  
EMBRYOS AND EARLY LARVAE.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=IA;  
CC IsoId=P10079-1; Sequence=Displayed;  
CC Name=IB;  
CC IsoId=P10079-2; Sequence=VSP 000451;  
CC -!- DEVELOPMENTAL STAGE: Moderate levels in unfertilized eggs and  
during early cleavage, then rapidly increases in abundance between  
late morula and mesenchyme blastula stages to maximal levels  
maintained through subsequent stages. Expressed both maternally  
and zygotically.  
CC -!- SIMILARITY: Contains 21 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 CUB domain.  
CC -!- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR TO  
AVIDIN/STREPTAVIDIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L08692; AAA62164.1; --  
DR EMBL; L08692; AAA62163.1; --  
DR EMBL; X17530; CAA35571.1; --  
DR EMBL; M17421; AAA30050.1; --  
DR EMBL; X17533; CAA35573.1; --  
DR PIR; A40136; A40136.  
DR HSSP; P01132; 1EGF.  
DR InterPro; IPR000152; Asx hydroxyl\_s.  
DR InterPro; IPR005469; Avidin.  
DR InterPro; IPR005468; Avidin/str.  
DR InterPro; IPR000859; CUB.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.

DR InterPro; IPR006209; EGF\_like.  
DR Pfam; PF01382; Avidin; 1.  
DR Pfam; PF00431; CUB; 1.  
DR Pfam; PF00008; EGF; 21.  
DR PRINTS; PR00709; AVIDIN.  
DR PRINTS; PR00010; EGF\_BLOOD.  
DR SMART; SM00042; CUB; 1.  
DR SMART; SM00179; EGF\_CA; 20.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 19.  
DR PROSITE; PS00577; AVIDIN; 1.  
DR PROSITE; PS01180; CUB; 1.  
DR PROSITE; PS00022; EGF\_1; 19.  
DR PROSITE; PS01186; EGF\_2; 19.  
DR PROSITE; PS50026; EGF\_3; 21.  
DR PROSITE; PS01187; EGF\_CA; 18.  
KW Biotin; Alternative splicing; EGF-like domain; Repeat; Signal;  
KW Glycoprotein; Calcium-binding.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1064 FIBROPELLIN I.  
FT DOMAIN 20 55 EGF-LIKE 1.  
FT DOMAIN 62 175 CUB.  
FT DOMAIN 176 212 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 214 250 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 252 288 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 290 326 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 328 364 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 366 402 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 404 440 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 442 478 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 480 516 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 518 554 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 556 592 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 594 630 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 632 668 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 670 706 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 708 744 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 746 782 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 784 820 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 822 858 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 860 896 EGF-LIKE 20.  
FT DOMAIN 898 934 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 936 1064 AVIDIN-LIKE.  
FT DISULFID 23 34 BY SIMILARITY.  
FT DISULFID 28 43 BY SIMILARITY.  
FT DISULFID 45 54 BY SIMILARITY.  
FT DISULFID 180 191 BY SIMILARITY.  
FT DISULFID 185 200 BY SIMILARITY.  
FT DISULFID 202 211 BY SIMILARITY.  
FT DISULFID 218 229 BY SIMILARITY.  
FT DISULFID 223 238 BY SIMILARITY.  
FT DISULFID 240 249 BY SIMILARITY.  
FT DISULFID 256 267 BY SIMILARITY.  
FT DISULFID 261 276 BY SIMILARITY.  
FT DISULFID 278 287 BY SIMILARITY.  
FT DISULFID 294 305 BY SIMILARITY.  
FT DISULFID 299 314 BY SIMILARITY.  
FT DISULFID 316 325 BY SIMILARITY.  
FT DISULFID 332 343 BY SIMILARITY.  
FT DISULFID 337 352 BY SIMILARITY.  
FT DISULFID 354 363 BY SIMILARITY.  
FT DISULFID 370 381 BY SIMILARITY.  
FT DISULFID 375 390 BY SIMILARITY.  
FT DISULFID 392 401 BY SIMILARITY.  
FT DISULFID 408 419 BY SIMILARITY.  
FT DISULFID 413 428 BY SIMILARITY.  
FT DISULFID 430 439 BY SIMILARITY.  
FT DISULFID 446 457 BY SIMILARITY.  
FT DISULFID 451 466 BY SIMILARITY.  
FT DISULFID 468 477 BY SIMILARITY.  
FT DISULFID 484 495 BY SIMILARITY.  
FT DISULFID 489 504 BY SIMILARITY.  
FT DISULFID 506 515 BY SIMILARITY.  
FT DISULFID 522 533 BY SIMILARITY.

FT	DISULFID	527	542	BY SIMILARITY.
FT	DISULFID	544	553	BY SIMILARITY.
FT	DISULFID	560	571	BY SIMILARITY.
FT	DISULFID	565	580	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
FT	DISULFID	598	609	BY SIMILARITY.
FT	DISULFID	603	618	BY SIMILARITY.
FT	DISULFID	620	629	BY SIMILARITY.
FT	DISULFID	636	647	BY SIMILARITY.
FT	DISULFID	641	656	BY SIMILARITY.
FT	DISULFID	658	667	BY SIMILARITY.
FT	DISULFID	674	685	BY SIMILARITY.
FT	DISULFID	679	694	BY SIMILARITY.
FT	DISULFID	696	705	BY SIMILARITY.
FT	DISULFID	712	723	BY SIMILARITY.
FT	DISULFID	717	732	BY SIMILARITY.
FT	DISULFID	734	743	BY SIMILARITY.
FT	DISULFID	750	761	BY SIMILARITY.
FT	DISULFID	755	770	BY SIMILARITY.
FT	DISULFID	772	781	BY SIMILARITY.
FT	DISULFID	788	799	BY SIMILARITY.
FT	DISULFID	793	808	BY SIMILARITY.
FT	DISULFID	810	819	BY SIMILARITY.
FT	DISULFID	826	837	BY SIMILARITY.
FT	DISULFID	831	846	BY SIMILARITY.
FT	DISULFID	848	857	BY SIMILARITY.
FT	DISULFID	864	875	BY SIMILARITY.
FT	DISULFID	869	884	BY SIMILARITY.
FT	DISULFID	886	895	BY SIMILARITY.
FT	DISULFID	902	913	BY SIMILARITY.
FT	DISULFID	907	922	BY SIMILARITY.
FT	DISULFID	924	933	BY SIMILARITY.
FT	CARBOHYD	30	30	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	851	851	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	477	780	Missing (in isoform IB). /FTID=VSP.000451.
FT	CONFLICT	279	279	L -> S (IN REF. 2).
SQ	SEQUENCE	1064	AA; 112072	MW; 2E569CA012ED6D09 CRC64;
Query Match 2.8%; Score 423.5; DB 1; Length 1064;				
Best Local Similarity 27.6%; Pred. No. 1.6e-16;				
Matches 131; Conservative 59; Mismatches 150; Indels 135; Gaps 30;				
QY	545	NDGKESEVVS-FLTTAIE-----SVDNCPSPN-CYNGDCI-----SGTCHCFLGLG 589		
Db	605	NGGACMDVVNGFVCTCLPGWEGTNCETDEACASSPCVNGGLCVQDVQNSYVCFCLPGFTG 664		
QY	590	PDCGR-----ASCPVLCSGNGQYM-----KGRCLCHSGWKGAECDDVPTNQCIDVACSNHGT 640		
Db	665	IHCGETEIDECASSP--CLNGGQCIDRVDSYECVCAAGYTAVRQCINIDECASAPCQNGGV 722		
QY	641	CITG-----TCICNPGYKGESCE-EVD-CMDPTCSGRGVCVR-----GECHCFVGMWGTFNCE 690		
Db	723	CVDGVNGYVNCAPGYTGDNCETEIDECASMPCLNGGACIEMVNGYTCCQVAGYTGVICE 782		
QY	691	TPRATCLDQCSG-----HGTFLPDT---GLCSDPSWTGHDG--SIEICAAD-CGGHGVGV 740		
Db	783	TD-----IDECASAPCQNGGVCTDTINGVICACVPFTGTSNCETNIDECASDPCLNGGICV 838		
QY	741	GG-----TCRCEDGMGAACD--QRACHPR-CAEHGTC----RDGKCECSPGWNGEHTCTIA 789		
Db	839	DGVNGFVCQCPNYSGTYTEISLDACRSMPCQNGATCVNVGADYVCEVCPYAGQNCBID 898		
QY	790	HYLDRVVKEGCPGL-CNGNGRCTLDLNGWHVCQLGWRGAGCDTSMETA-CGDSKDNDDG 847		
Db	899	-----INE-CASLPCQNGGLCIDGIAGYTCQRLGYIGNVCE---EVGFC----DLEGM 944		
QY	848	GLVDCMDPDCCQLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKFLVGRDST 907		
Db	945	WYNECND-----QVTIT-----KTST 960		
QY	908	HIIPGE-----NPFDDGGHACVIRGVQVMTSDGTPLVGVNISFVNNPLFGYTISRQDG 958		

Db	961	GMMLGDYMTYNERALGYAA-----PTVVVGYASNNYDFPSFGFTVVRDNG 1005		
RESULT 13				
NTC1_RAT				
ID	NTC1_RAT	STANDARD;	PRT;	2531 AA.
AC	Q07008;			
DT	01-NOV-1995	(Rel. 32, Created)		
DT	15-JUL-1999	(Rel. 38, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Neurogenic locus notch homolog protein 1 precursor (Notch 1).			
GN	NOTCH1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Schwann cell;			
RX	MEDLINE=92111383; PubMed=1764995;			
RA	Weinmaster G., Roberts V.J., Lemke G.;			
RT	"A homolog of Drosophila Notch expressed during mammalian			
RT	development.";			
RL	Development 113:199-205(1991).			
RN	[2]			
RP	REVISIONS TO 1652-1653.			
RA	Weinmaster G.;			
RL	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	FUNCTION.			
RX	MEDLINE=21094508; PubMed=11182080;			
RA	Tanigaki K., Nogaki F., Takahashi J., Tashiro K., Kurooka H.,			
RA	Honjo T.;			
RT	"Notch1 and Notch3 instructively restrict bFGF-responsive multipotent			
RT	neural progenitor cells to an astroglial fate.";			
RL	Neuron 29:45-55(2001).			
RN	[4]			
RP	TISSUE SPECIFICITY.			
RX	MEDLINE=93202015; PubMed=1295745;			
RA	Weinmaster G., Roberts V.J., Lemke G.;			
RT	"Notch2: a second mammalian Notch gene.";			
RL	Development 116:931-941(1992).			
RN	[5]			
RP	TISSUE SPECIFICITY.			
RX	MEDLINE=21331789; PubMed=11438922;			
RA	Irvin D.K., Zurcher S.D., Nguyen T., Weinmaster G., Kornblum H.I.;			
RT	"Expression patterns of Notch1, Notch2, and Notch3 suggest multiple			
RT	functional roles for the Notch-DSL signaling system during brain			
RT	development.";			
RL	J. Comp. Neurol. 436:167-181(2001).			
CC	-!- FUNCTION: Functions as a receptor for membrane-bound ligands			
CC	Jagged1, Jagged2 and Deltal to regulate cell-fate determination.			
CC	Upon ligand activation through the released notch intracellular			
CC	domain (NICD) it forms a transcriptional activator complex with			
CC	RBP-J kappa and activates genes of the enhancer of split locus.			
CC	Affects the implementation of differentiation, proliferation and			
CC	apoptotic programs (By similarity). Acts instructively to control			
CC	the cell fate determination of CNS multipotent progenitor cells,			
CC	resulting in astroglial induction and neuron/oligodendrocyte			
CC	suppression.			
CC	-!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-			
CC	terminal fragment N(EC) which are probably linked by disulfide			
CC	bonds (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein. Following			
CC	proteolytical processing NICD is translocated to the nucleus (By			
CC	similarity).			
CC	-!- TISSUE SPECIFICITY: Expressed in the brain, kidney and spleen.			
CC	Expressed in postnatal central nervous system (CNS) germinal zones			
CC	and, in early postnatal life, within numerous cells throughout the			
CC	CNS. Found in both subventricular and ventricular germinal zones.			
CC	-!- DEVELOPMENTAL STAGE: In the embryo, highest levels occur between			
CC	days 12 and 14 and decrease rapidly to much lower levels in the			









QY	942	-----SFVN-----	NPLFGVTISRQDGSFDLVT--	NGGISILRFR	977
Db	669	NDCSQNPCHYGTCRDLVNDFYCDCKNGWKGTCHSRDQCDEATCNNG	-----	718	
QY	978	PFITQHTLWLPDRFFVMETIIMRHEENEIPSC-----	DLSNFAR-----	PNP-----	1021
Db	719	-----TCYDEVDTFKM-----	CPGWEGTTCNIARNSSCLPNPCHNGG	757	
QY	1022	--VWSPSLTSSFASCAE--	KGPIVPEIQALQEEHISISGCKMRLSYLSSR-----	1067	
Db	758	TCVWNGD--SFTCVCKEGWEGPICTQ-----	NTNDCSPHCYNSGTCVDDGNWYRC	806	
QY	1068	--TPGYKSVLRISLTHPTIPFNLMKVHLMVAVEGRLFRKWFAAPDL	SYFIWDKTDVYN	1125	
Db	807	ECAPGFAG-----	PDCRININECQSSPCAFA-----	TCVDEINGY-----	842
QY	1126	QKVFLSEAFVSUGVEYESCPDLILWEKRTTVLQGYEI-DASKLGGWSLDKXHALNIQSG	1184		
Db	843	-----QCICPPGHSGAKCHEV--	SGRSCITMGRVILDGAK--	WDDDCNTCQCLNGR	889
QY	1185	I-----	LHKNGENQFVSQQPPVIGS-----	IMNGRRRRSISCPSCNGLAD	1225
Db	890	VACSKVWCGRPRLHKGHGECPNGQSCIPVLDDQCFVRPCTGAGECRSSS	-----	940	
QY	1226	GNKLLAPVALTCGSDGSLYVD-----	FNYIRRIFFPSGNVTNIL--	ELRNKDFRHSHP	1277
Db	941	-----LQPVKTCTSDS--	YYQDNCANITFTFNKEMMSPGLTTEHICSELRLNLILKNVSA	994	
QY	1278	AHKYYLATDPMGAVFLSDNSRRVPKIKSTVVVDLVKNSEVAGTGDQCLFDDTRCG	1337		
Db	995	EYSIYIACEPS-----	LSANN-----	EIHVAISAEDIRDDGNPVEITDKIIDLVSKR--	1042
QY	1338	DGGKATEATLTNPR	1351		
Db	1043	DGNSSLIAAAVEVR	1056		

Search completed: June 24, 2004, 16:19:49  
Job time : 31 secs

**This Page Blank (uspto)**